

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ALGORITMA *STEMMING* TEKS BAHASA MINANG BERBASIS ATURAN TATA BAHASA

TUGAS AKHIR

Diajukan Sebagai Salah Satu Syarat
Untuk Memperoleh Gelar Sarjana Teknik
Pada Jurusan Teknik Informatika

Oleh

DEWITA PUTRI

11351201047



RIAU

**FAKULTAS SAINS DAN TEKNOLOGI
UNIVERSITAS ISLAM NEGERI SULTAN SYARIF KASIM RIAU**

2021

LEMBAR PERSETUJUAN

ALGORITMA *STEMMING* TEKS BAHASA MINANG

BERBASIS ATURAN TATA BAHASA

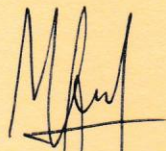
TUGAS AKHIR

Oleh

DEWITA PUTRI
11351201047

Telah diperiksa dan disetujui sebagai Laporan Tugas Akhir
di Pekanbaru, pada tanggal 05 Februari 2021

Pembimbing I,



Muhammad Fikry, S.T., M.Sc
NIP. 19801018 200710 1 002

Pembimbing II,



Yusra, S.T., M.T
NIP. 19840123 201503 2 001

LEMBAR PENGESAHAN

ALGORITMA *STEMMING* TEKS BAHASA MINANG BERBASIS ATURAN TATA BAHASA

TUGAS AKHIR

Oleh

DEWITA PUTRI
11351201047


Telah dipertahankan di depan sidang dewan penguji
sebagai salah satu syarat untuk memperoleh gelar Sarjana Teknik
Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau
di Pekanbaru, pada tanggal 05 Februari 2021

Pekanbaru, 05 Februari 2021

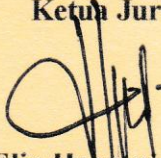
Mengesahkan,



Dekan


Dr. Drs. Ahmad Darmawi, M.Ag.
NIP. 19660604 199203 1 004

Ketua Jurusan


Dr. Elin Haerani, S.T., M.Kom.
NIP. 19810523 200710 2 003

DEWAN PENGUJI

Ketua : Novriyanto, S.T, M.Sc

Sekretaris : Muhammad Fikry, S.T., M.Sc

Anggota I : Yusra, S.T., M.T.

Anggota II : Febi Yanto, M.Kom

Anggota III: Elvia Budianita, S.T., M.Cs

LEMBAR HAK KEKAYAAN INTELEKTUAL

Tugas Akhir yang tidak diterbitkan ini terdaftar dan tersedia di Perpustakaan Universitas Islam Negeri Sultan Syarif Kasim Riau adalah terbuka untuk umum dengan ketentuan bahwa hak cipta pada penulis. Referensi kepustakaan diperkenankan dicatat, tetapi pengutipan atau ringkasan hanya dapat dilakukan seizin penulis dan harus disertai dengan kebiasaan ilmiah untuk menyebutkan sumbernya.

Penggandaan atau penerbitan sebagian atau seluruh Tugas Akhir ini harus memperoleh izin dari Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau. Perpustakaan yang meminjamkan Tugas Akhir ini untuk anggotanya diharapkan untuk mengisi nama, tanda peminjaman dan tanggal pinjam pada form peminjaman.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERNYATAAN

Dengan ini saya menyatakan bahwa dalam Tugas Akhir ini tidak terdapat karya yang pernah diajukan untuk memperoleh gelar kesarjanaan di suatu Perguruan Tinggi, dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan didalam daftar pustaka.

Pekanbaru, 05 Februari 2021

Yang membuat pernyataan,

DEWITA PUTRI
11351201047

UIN SUSKA RIAU



Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LEMBAR PERSEMBAHAN

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

*Alhamdulillahirrabbi' alamin
Segala Puji bagiMu Tuhan Semesta Alam
Alhamdulillah aku ucapkan
Atas berkah dan rahmat yang Engkau berikan
Aku dapat berjuang dalam menyelesaikan Tugas Akhir ini*

*Ya Allah
Ampunilah dosa ibu dan ayah ku.....
Kasihilah mereka seperti mereka mengasihiku di waktu aku kecil dan dewasa ini Berikanlah
kebahagiaan kepada mereka ya Allah. Aamiin*

*Untuk Ayah (Mofrizal) dan Ibu (Zulbaida)
Terimakasih Adek ucapkan atas semua yang telah Ayah dan Ibu berikan, maafkan selama ini
adek selalu membuat Ayah dan Ibu susah, Tugas Akhir ini adek persembahkan untuk Ayah dan
Ibu yang tidak pernah henti-hentinya mendoakan anaknya untuk menjadi orang yang sukses baik
di dunia maupun di akhirat kelak. Aamiin. Semoga dengan hasil dari perjuangan adek ini, Ayah
dan Ibu bisa bangga dengan adek. Jasa Ayah dan Ibu tidak akan pernah terbalaskan. Adek akan
berusaha untuk selalu menorehkan senyuman dan air mata bahagia diwajah Ayah dan Ibu.
Semoga Allah SWT selalu melimpahkan kebahagiaan dan kesehatan untuk Ayah dan Ibu. Aamiin
ya Rabbal' alamin.*

*Untuk kakakku (Diana Wahyuni,S.Kom) dan kedua adikku (Dea Adilla dan Diah Syifa Aprilia)
Terimakasih ya selama ini kalian selalu memberikan dukungan dan selalu membantu adek dalam
menyelesaikan Tugas Akhir ini. Untuk dea dan syifa, kakak berharap semoga kalian bisa
menyelesaikan sekolahnya kejenjang pendidikan yang lebih tinggi, Aamiin.*

*Terimakasih untuk Dosen pembimbing dan para dosen yang telah membantu Dewita dalam
menyelesaikan Tugas Akhir ini.*

*Dan terimakasih juga untuk teman-teman TIF'G yang telah membantu dalam menyelesaikan
Tugas Akhir ini, yang namanya tidak bisa disebutkan satu persatu, semoga kebaikan kalian
semua dilimpahkan Pahala oleh ALLAH SWT. Aamiin.*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

ALGORITMA *STEMMING* TEKS BAHASA MINANG BERBASIS ATURAN TATA BAHASA

DEWITA PUTRI
11351201047

Tanggal Sidang : 05 Februari 2021

Jurusan Teknik Informatika
Fakultas Sains dan Teknologi
Universitas Islam Negeri Sultan Syarif Kasim Riau

ABSTRAK

Dalam bahasa Minang terdapat kata dasar dan kata berimbuhan. Kata berimbuhan terbentuk dengan adanya penambahan imbuhan awalan, akhiran dan gabungan awalan akhiran sehingga kata tersebut menjadi bervariasi. Pada penelitian ini data uji yang digunakan berupa kata berimbuhan yang didapat dari lirik lagu bahasa Minang. Untuk menghapus imbuhan yang melekat pada kata dasar dilakukan proses *stemming*. *Stemming* merupakan proses penguraian yang digunakan untuk menemukan kata dasar dari kata yang mengalami imbuhan, dengan cara menghilangkan atau menghapus imbuhan yang melekat pada kata dasar. Teknik *stemming* dibagi menjadi beberapa tahapan, tahap pertama yaitu mengumpulkan aturan-aturan yang digunakan bahasa Minang, tahap kedua yaitu mengumpulkan kata dasar yang akan menjadi kamus basis data dan tahap ketiga yaitu penghapusan kata berimbuhan yang melekat pada kata dasar. Berdasarkan hasil pengujian yang dilakukan dari algoritma *stemming* yang telah dibuat, didapatkan nilai akurasi sebesar 100% dari 345 kata uji.

Kata Kunci: Algoritma, Bahasa Minang dan *Stemming*

MINANGNESE LANGUAGE STEMMING TEXT ALGORITHM BASED ON LANGUAGE RULE BASED

DEWITA PUTRI

11351201047

Session Date : 05 Februari 2021

Informatics Engineering Department

Faculty of Science and Technology

Sultan Syarif Kasim State Islamic University Riau

ABSTRACT

Minangnese language contains of words with affix and root words. The words with affix is a result from addition prefix at the beginning of the word, suffix at the end of the word and combine of prefix and suffix that makes it diverse. In this research the data that will be tested are words with affix collected from song with Minangnese Language. Stemming process is required to remove affix that attached to the root word. Stemming is a process to decomposite words with affix to find its root word by removing or clearing the affix. Steaming technic devide into several step, first step is collecting the rules that is used in Minangnese Language. Second step is to collect the root word that will be use as database dictionary and the third step is to remove the affix that is attached to the root word. The result of the test of this stemming algorithm is 100% accuracy of 345 words test.

Keyword: Algorithm, Minangnese Language and Stemming

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

KATA PENGANTAR

Assalamu'alaikum wa rahmatullahi wa barakatuh.

Alhamdulillahirabbil'alamin, Segala puji hanya bagi Allah SWT, karena berkat limpahan rahmat, karunia dan hidayah-Nya sehingga penulis mampu menyelesaikan penelitian dan penulisan laporan Tugas Akhir ini. Shalawat beriring salam tak lupa penulis ucapkan untuk junjungan kita Rasulullah Muhammad SAW, dengan mengucapkan *Allahumma Sholli'Alaa Sayyidina Muhammad Wa'Alaa Ali Sayyidina Muhammad* karena berkat jasa beliau kita bisa menikmati zaman yang penuh dengan ilmu pengetahuan seperti saat sekarang ini.

Tugas Akhir yang berjudul "**Algoritma Stemming Teks Bahasa Minang Berbasis Aturan Tata Bahasa**" ini disusun sebagai salah satu syarat untuk memperoleh gelar sarjana strata satu pada Program Studi Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau. Dalam pelaksanaan dan penyusunan Tugas Akhir ini, penulis banyak mendapatkan bimbingan, pengetahuan, pengalaman, motivasi dan dukungan dari berbagai pihak yang telah membantu hingga Tugas Akhir ini dapat diselesaikan secara baik. Untuk itu penulis ingin menyampaikan ucapan terima kasih yang sebesar-besarnya kepada:

1. Bapak Prof. Dr. Suyitno, M.Ag, selaku Rektor Universitas Islam Negeri Sultan Syarif Kasim Riau.
2. Bapak Dr. Ahmad Darmawi, M.Ag, selaku Dekan Fakultas Sains dan Teknologi Universitas Islam Negeri Sultan Syarif Kasim Riau.
3. Ibu Dr. Elin Haerani, S.T, M.Kom, selaku Ketua Jurusan Teknik Informatika Universitas Islam Negeri Sultan Syarif Kasim Riau.
4. Bapak Muhammad Irsyad S.T, M.T, selaku Pembimbing Akademik penulis selama menjalani perkuliahan di Jurusan Teknik Informatika.
5. Bapak Muhammad Fikry, S.T, M.Sc, selaku dosen pembimbing I Tugas Akhir penulis yang telah meluangkan waktunya, memberikan wawasan, motivasi

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

dan ilmu yang bermanfaat bagi penulis sehingga penulis dapat menyelesaikan Tugas Akhir ini dengan baik.

6. Ibu Yusra, S.T, M.T, selaku dosen pembimbing II, yang telah meluangkan waktunya dan banyak memberikan masukan kepada penulis dalam penyusunan tugas akhir ini.
7. Bapak Febi Yanto, M.Kom, selaku penguji I yang membantu memberikan masukan kepada penulis dalam penyusunan tugas akhir ini.
8. Ibu Elvia Budianita, S.T, M.Cs, selaku dosen penguji II, yang telah banyak meluangkan waktu, memberikan kritik dan saran dalam penulisan dan perbaikan laporan Tugas Akhir ini.
9. Bapak/Ibu Dosen Teknik Informatika yang selama ini memberikan ilmu kepada penulis dan juga nasihat-nasihat yang akan selalu diingat.
10. Kedua orang tua yang telah berjuang untuk memberikan pendidikan terbaik, Ayah Mofrizal dan Ibu Zulbaida. Terima kasih atas do'a dan pengorbanan yang Ayah dan Ibu beri demi kesuksesan anak-anaknya serta menjadi inspirasi, sumber semangat dan motivasi hidup dalam setiap langkah penulis untuk menyelesaikan laporan Tugas Akhir ini. Semoga beliau selalu dalam lindungan Allah SWT. *Aamiin Ya Rabbal'alamiin.*
11. Kakak saya Diana Wahyuni S.Kom dan kedua adik saya Dea Adilla dan Diah Syifa Aprilia, yang tak henti-hentinya menyemangati penulis selama mengerjakan Tugas Akhir ini.
12. Teman sekelas "Broken G" TIF G 2013 yang telah memberikan dukungan. Terima kasih telah berbagi ilmu dan cerita selama di bangku perkuliahan. Semoga kelak kita semua menjadi orang yang berguna bagi nusa dan bangsa serta bahagia di dunia dan di akhirat.
13. Teman-teman TIF angkatan 2013 serta junior TIF yang tidak mungkin disebutkan satu persatu.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

14. Semua pihak yang terlibat langsung maupun tidak langsung dalam pelaksanaan tugas akhir ini yang tidak dapat disebutkan satu persatu.

Semoga laporan ini dapat bermanfaat bagi penulis dan juga para pembaca laporan ini. Penulis berharap ada kritik dan saran dari pembaca laporan ini guna memperbaiki dan pengembangan dari laporan ini kedepannya. Kritik dan saran tersebut dapat dikirim ke email penulis yaitu dewita.putri@students.uin-suska.ac.id
Akhir kata penulis ucapkan terima kasih dan selamat membaca.

Wassalamu'alaikum wa rahmatullahi wa barakatuh.

Pekanbaru, 05 Februari 2021

DEWITA PUTRI

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR ISI

	Halaman
LEMBAR PERSETUJUAN	ii
LEMBAR PENGESAHAN.....	iii
LEMBAR HAK KEKAYAAN INTELEKTUAL	iv
LEMBAR PERNYATAAN	v
LEMBAR PERSEMBAHAN	vi
ABSTRAK	vii
ABSTRACT	viii
KATA PENGANTAR.....	ix
DAFTAR ISI.....	xii
DAFTAR GAMBAR.....	xv
DAFTAR TABEL	xvi
DAFTAR SIMBOL.....	xvii
DAFTAR LAMPIRAN	xviii
BAB I PENDAHULUAN.....	I-1
1.1. Latar Belakang.....	I-1
1.2. Rumusan Masalah.....	I-4
1.3. Batasan Masalah	I-4
1.4. Tujuan	I-5
1.5. Sistematika Penulisan	I-5
BAB II LANDASAN TEORI.....	II-1
2.1. Natural Language Prosessing (NLP)	II-1

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.2.	Bahasa Minang	II-3
2.3.	Pola Bahasa Minangkabau	II-5
2.4.	Morfologi Bahasa Minang	II-11
2.4.1.	Imbuhan dan Distribusinya	II-11
2.4.2.	Morfem	II-15
2.4.3.	Klasifikasi Kata	II-16
2.4.4.	Proses Morfologis.....	II-17
2.5.	<i>Stemming</i>	II-18
2.6.	Penelitian Terkait.....	II-19

BAB III METODOLOGI PENELITIANIII-1

3.1.	Perumusan Masalah	III-2
3.2.	Pengumpulan Data.....	III-2
3.3.	Analisa dan Perancangan Algoritma <i>Stemming</i>	III-2
3.3.1.	Analisa Kebutuhan Data	III-2
3.3.2.	Analisa Proses <i>Stemming</i>	III-2
3.3.3	Perancangan.....	III-3
3.4.	Implementasi dan Pengujian	III-3
3.4.1.	Implementasi.....	III-3
3.4.2.	Pengujian.....	III-4
3.5.	Kesimpulan dan Saran	III-4

BAB IV ANALISA DAN PERANCANGANIV-1

4.1.	Analisa Data	IV-1
4.1.1.	Pengumpulan Data.....	IV-2

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4.1.2.	Analisa Aturan Pembentukan Imbuhan	IV-3
4.1.3.	Aturan Penghapusan Imbuhan	IV-5
4.2.	Perancangan.....	IV-7
4.2.1.	Perancangan <i>Flowchart</i>	IV-7
4.2.2.	Perancangan <i>Pseudocode</i>	IV-9
4.2.3.	Perancangan Basis Data (<i>Database</i>).....	IV-9
4.2.4.	Perancangan Antar Muka (<i>Interface</i>)	IV-9
BAB V IMPLEMENTASI DAN PENGUJIAN		V-1
5.1.	Implementasi Penerapan Algoritma	V-1
5.1.1.	Lingkungan Implementasi	V-1
5.1.2.	Batas Lingkungan Implementasi	V-2
5.2.	Implementasi Antar Muka.....	V-2
5.2.1.	Tampilan Input Kata Berimbuhan	V-2
5.2.2.	Tampilan Output Kata Dasar	V-3
5.2.3.	Tampilan Akurasi Kata Uji	V-3
5.3.	Pengujian Algoritma <i>stemming</i> Bahasa Minang	V-4
5.4.	Analisa Hasil Pengujian	V-14
BAB VI PENUTUP		VI-1
6.1.	Kesimpulan.....	VI-1
6.2.	Saran	VI-1
DAFTAR PUSTAKA.....		xix

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR GAMBAR

Gambar	Halaman
2.1 Cover Buku Kamus Bahasa Minangkabau-Indonesia	II-5
2.2 Cover Buku Tata Bahasa Minangkabau.....	II-6
3.1 Tahap Metode Penelitian.....	III-1
4.1 Gambaran Umum Proses Stemming	IV-1
4.2 Flowchart Algoritma Stemming Bahasa Minang	IV-8
4.3 Rancangan Tampilan Input Kata Berimbuhan	IV-10
4.4 Rancangan Tampilan Output Kata Dasar.....	IV-10
4.5 Rancangan Tampilan Kata Uji Stemming Bahasa Minang.....	IV-11
5.1 Tampilan Input Kata Berimbuhan Bahasa Minang	V-2
5.2 Tampilan Output Kata Dasar Bahasa Minang.....	V-3
5.3 Tampilan Akurasi Kata Uji Bahasa Minang	V-4

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR TABEL

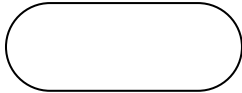
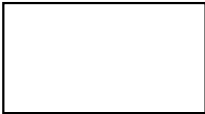

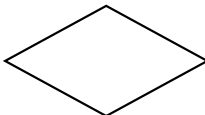
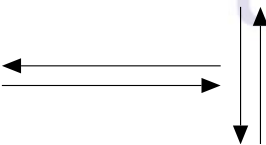

Tabel	Halaman
2.1 Reduplikasi dan Imbuhan.....	II-18
2.2 Penelitian Terkait.....	II-19
4.1 Jenis-jenis imbuhan bahasa minang.....	IV-2
4.2 Aturan Pembentukan Imbuhan Awalan Bahasa Minang	IV-3
4.3 Aturan Pembentukan Imbuhan Akhiran Bahasa Minang.....	IV-4
4.4 Aturan Pembentukan Imbuhan Awalan Akhiran Bahasa Minang.....	IV-5
4.5 Aturan Penghapusan Imbuhan Awalan dalam bahasa minang.....	IV-6
4.6 Struktur Tabel Kata Dasar.....	IV-9
4.7 Struktur Tabel Kata Uji.....	IV-9
5.1 Kata Uji	V-5

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR SIMBOL

Keterangan notasi simbol *Flowchart* :

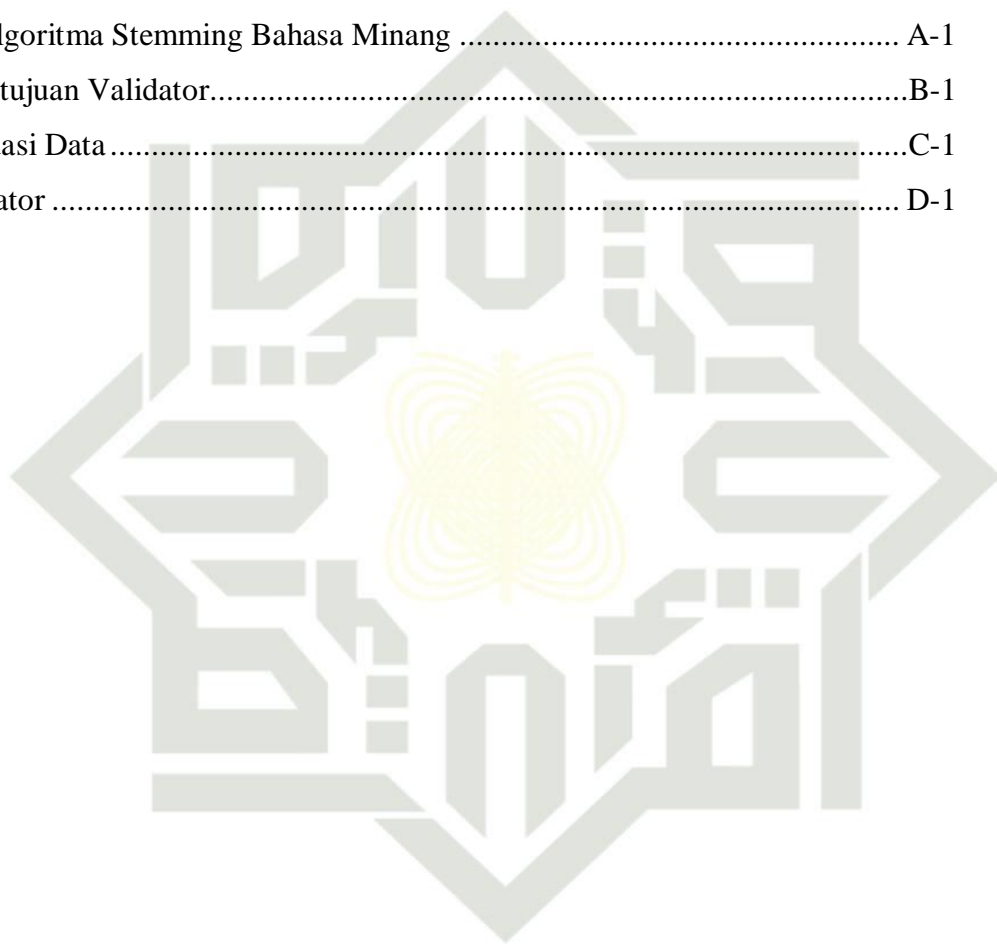
Simbol	Keterangan
	Terminator: Simbol ini menunjukkan awal dan akhir program.
	Proses: digambarkan dengan persegi panjang. Umumnya mendefinisikan mekanisme proses ataupun kegiatan yang sedang dilakukan oleh komputer.
	Dokumen: digambarkan dengan kombinasi persegi panjang dan garis lengkung. Umumnya mendefinisikan dokumen masukan (formulir) dan dokumen keluaran (laporan) atau dikatakan dokumen input dan output.
	Verifikasi: digambarkan dengan belah ketupat. Symbol ini digunakan untuk memutuskan apakah valid atau tidak validnya suatu program.
	Arus Data: Menunjukkan arus dari proses yang sedang berjalan.
	Data: Simbol yang digunakan untuk mendeskripsikan data <i>input</i> atau <i>output</i> yang digunakan.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR LAMPIRAN

Lampiran	Halaman
Kata Uji Algoritma Stemming Bahasa Minang	A-1
Surat Persetujuan Validator.....	B-1
Surat Validasi Data	C-1
Data Validator	D-1



UIN SUSKA RIAU

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB I

PENDAHULUAN

1.1. Latar Belakang

Bahasa merupakan alat komunikasi masyarakat yang terdiri dari sistem lambang bunyi dan digunakan bersama secara konvensional untuk berkomunikasi dan berinteraksi sesama manusia (Arifin, 1990). Bahasa sehari-hari merupakan bahasa yang dipakai sehari-hari dalam berkomunikasi antar sesama anggota keluarga, tetangga dan masyarakat sekitar. Bahasa sehari-hari yang dipakai terbentuk karena proses interaksi sosial (BPS, 2010).

Dalam Sensus Penduduk tahun 2010, bahasa sehari-hari dibagi menjadi tiga kelompok bahasa, yaitu:

1. Bahasa Indonesia
2. Bahasa Daerah
3. Bahasa Asing

Bahasa daerah merupakan bahasa yang digunakan oleh penduduk asli (suku bangsa) suatu daerah dengan ciri khas daerah yang bersangkutan. Contoh: Bahasa Minangkabau, Bahasa Batak, Bahasa Sunda (Arifin, 1990). Bahasa Minangkabau merupakan sarana komunikasi bagi masyarakat minang (Sumatra Barat). Bahasa ini memiliki berbagai sebutan, yaitu bahasa Minang atau bahasa Padang (Grimes, dalam Nadra, 2006) dan bahasa Melayu Minangkabau (Muhajir, dalam Nadra, 2006). Jumlah penutur bahasa Minangkabau menduduki peringkat kelima dari sepuluh bahasa daerah terbesar yang ada di Indonesia. Di Sumatra, bahasa Minangkabau adalah bahasa kedua terbesar setelah bahasa Melayu (Burhanuddin et al., 2012).

Sumatra Barat yang merupakan daerah Asal Minangkabau, terbagi atas dua bagian yaitu *darek* dan *rantau*. *Darek* adalah daerah tua Minangkabau, yang terdiri atas tiga *luhak*, yaitu Luhak Tanah datar, Luhak Agam, dan Luhak Limo Puluah Koto. Ketiga daerah tersebut terletak di sekitar Gunung Marapi, Singgalang, dan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Sago. Daerah *rantau* meliputi pesisir pantai Sumatra, di antaranya Padang, Pariaman, Pesisir Selatan, Kerinci, Natal, dan Indera Giri (Burhanuddin et al., 2012).

Seiring dengan perkembangan ilmu pengetahuan dan teknologi, bahasa Minangkabau terus berubah sebagaimana halnya bahasa-bahasa daerah lain yang ada di Indonesia. Bahasa Indonesia dan juga bahasa asing lainnya turut mempengaruhi penutur bahasa tersebut. Media audio visual, seperti radio, televisi dan *youtube* memiliki andil dalam perubahan tersebut. Melalui media inilah semua kosakata dan istilah diperkenalkan kepada generasi kita. Sebagaimana kita ketahui, bahasa Minangkabau memiliki keterbatasan dalam mengodifikasi pengetahuan *modern*. Oleh karena itu, ruang pemakaian bahasa tersebut makin menyempit.

Selain itu, secara tidak langsung, para orang tua (keluarga muda) juga turut menjauhkan generasinya dengan bahasa ibunya sendiri. Hal itu ditunjukkan dengan bahasa yang digunakan dalam berkomunikasi sehari-hari. Pada saat ini, mereka telah membiasakan anak-anak mereka menggunakan bahasa Indonesia sejak dini, walau terkadang bahasanya “*sakarek ula, sakarek baluik*” (tidak sepenuhnya berbahasa Indonesia). Jadi, tidaklah mengherankan kalau saat ini generasi muda mulai terbata-bata dengan bahasa ibunya sendiri (Burhanuddin et al., 2012).

Seiring dengan berkurangnya penutur yang menggunakan dan memahami bahasa Minangkabau, lambat laun makin banyak pula kosakata, mamangan, pepatah-petitih, ungkapan, pemeo bahasa Minangkabau yang tidak terpakai lagi dan berangsur lenyap. Oleh karena itu, kekayaan bahasa Minangkabau ini harus diselamatkan dan dilestarikan dalam bentuk digital agar kalangan muda pada masa ini dapat mengenal sejarah dan kebudayaan bahasa Minangkabau.

Stemming merupakan suatu proses untuk menemukan kata dasar dari sebuah kata. *Stemming* merupakan proses yang memetakan bentuk varian kata menjadi kata dasarnya (Tala, 2004). Untuk *stemming* dalam bahasa Indonesia, terdapat dua jenis metode *stemming* yang sudah ada, yaitu algoritma *stemming* yang berbasis kamus

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

(*dictionary based*) dan algoritma *stemming* yang berbasis non-kamus (*purely rule based*) (Manase Sahat H Simarangkir, 2017).

Beberapa penelitian terkait mengenai *stemming* bahasa daerah antara lain, penelitian yang dilakukan oleh Purwoko dan Pascasarjana pada tahun 2011 dengan judul model *stemming* berbasis kamus untuk dokumen berbahasa sunda yang menghasilkan algoritma *stemming* dengan memberikan hasil kata sebesar 91,38% dan mengurangi jumlah kata pada indeks sebesar 49,31%. Kemudian penelitian yang dilakukan oleh Mar'atus Madia ditahun 2016 dengan judul *stemming* bahasa jawa untuk mencari akar kata dalam bahasa jawa dengan aturan analisis kontrasif afiksasi verba akurasi yang didapatkan dari hasil pengujian algoritma *stemming* sebesar 92,881% dari 16.745 kata bahasa Jawa.

Penelitian terkait lainnya oleh Ledy Agusta yang berjudul perbandingan algoritma *stemming* porter dengan algoritma nazief & adriani untuk *stemming* dokumen teks bahasa indonesia pada tahun 2009, penelitian yang dilakukan oleh Ananda Wulandari, Kemas Rahmat S.W, Ade Romadhony yang berjudul *pattern-based stemmer analysis and implementation on arabic text* pada tahun 2011, penelitian yang dilakukan oleh Lasmedi Afuan yang berjudul *stemming dokumen teks bahasa indonesia menggunakan algoritma porter* pada tahun 2013, penelitian yang dilakukan oleh Damar Aji Asmara, Dewi Khairani, Siti Ummi Masruroh yang berjudul *penerapan algoritma paice atau husk untuk stemming pada kamus bahasa inggris ke bahasa indonesia* pada tahun 2013, penelitian yang dilakukan oleh Noverdy Anggara, Ade Romadhony, Mahmud Dwi Suliiyo yang berjudul *implementasi modifikasi algoritma enhanced confix stripping stemmer pada teks bahasa indonesia* pada tahun 2013, penelitian yang dilakukan oleh Rakhmad Maulidi yang berjudul *stemmer untuk bahasa madura dengan modifikasi metode enhanced confix stripping stemmer* pada tahun 2016, penelitian yang dilakukan oleh Julianto Wibowo yang berjudul *aplikasi penentuan kata dasar dari kata berimbuhan pada kalimat bahasa indonesia dengan algoritma stemming* pada tahun 2016, penelitian yang dilakukan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

oleh Dian Novitasari yang berjudul perbandingan algoritma stemming porter dengan arifin setiono untuk menentukan tingkat ketepatan kata dasar pada tahun 2016, penelitian yang dilakukan oleh Gusti Ngurah Mega Nata, Putu Pande Yudiastra yang berjudul *stemming* teks *sor-singgih* bahasa bali pada tahun 2017, penelitian yang dilakukan oleh Dwi Wahyudi, Teguh Susyanto, Didik Nugroho yang berjudul implementasi dan analisis algoritma stemming nazief & adriani dan porter pada dokumen berbahasa indonesia pada tahun 2017, penelitian yang dilakukan oleh Manase Sahat H Simarankir yang berjudul studi perbandingan algoritma - algoritma *stemming* untuk dokumen teks bahasa indonesia pada tahun 2017, dan penelitian yang dilakukan oleh Adhi Prasadhatama yang berjudul analisis perbandingan algoritma nazief & adriani dengan algoritma idris untuk pencarian kata dasar pada teks bahasa indonesia berbasis *website* pada tahun 2017.

Sejauh ini belum ada penelitian tentang algoritma *stemming* bahasa Minang. Dari penjelasan latar belakang di atas maka penulis melakukan penelitian tugas akhir, algoritma *stemming* teks bahasa Minang berbasis aturan tata bahasa.

1.2. Rumusan Masalah

Berdasarkan latar belakang di atas, rumusan masalah yang dapat diambil yaitu bagaimana membuat algoritma *stemming* bahasa minang berbasis aturan tata bahasa.

1.3. Batasan Masalah

Supaya pembahasan penelitian ini tidak menyimpang dari apa yang telah dirumuskan, maka diperlukan batasan-batasan. Batasan-batasan dalam penelitian ini yaitu:

1. Kata-kata dasar bersumber dari kamus bahasa Minangkabau - Indonesia (Rusmali, M dkk., 1985).
2. Morfologi bersumber dari buku Tata Bahasa Minangkabau. (Ayub, Asni dkk., 1993).
3. Teks yang diambil bersumber dari lirik lagu bahasa Minangkabau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1.4. Tujuan

Adapun tujuan dari penelitian ini sebagai berikut:

1. Membuat algoritma penentuan kata dasar untuk bahasa daerah minang.
2. Menguji akurasi dari algoritma tersebut.

1.5. Sistematika Penulisan

Sistematika penulisan Tugas Akhir ini diatur sedemikian rupa sehingga segala kebutuhan yang dipergunakan di dalam penelitian dapat dipahami dengan mudah.

Sistematika penulisan laporan Tugas Akhir ini adalah sebagai berikut:

BAB I PENDAHULUAN

Bab ini menjelaskan latar belakang, rumusan masalah, batasan masalah, tujuan penelitian dan sistematika penulisan laporan.

BAB II LANDASAN TEORI

Bab ini menjelaskan dasar-dasar teori yang digunakan sebagai landasan dalam pembuatan Tugas Akhir seperti: Pengertian NLP, Pengertian Bahasa Minang, Pengertian *Stemming* dan Penelitian Terkait.

BAB III METODOLOGI PENELITIAN

Bab ini menjelaskan metodologi yang digunakan dalam penelitian yaitu: Perumusan Masalah, Pengumpulan Data, Perancangan, Implementasi, Pengujian, Kesimpulan dan Saran.

BAB IV ANALISA DAN PERANCANGAN

Bab ini membahas analisa langkah kerja dan perancangan algoritma yang akan dibuat sesuai dengan metodologi yang digunakan.

BAB V IMPLEMENTASI DAN PENGUJIAN

Pada bab ini akan dibahas implementasi dan pengujian hasil penelitian sesuai yang tertera pada rumusan masalah dan pembahasan.

BAB VI PENUTUP

Penutup merupakan jawaban atas rumusan masalah dalam penelitian dan juga intisari dari penelitian serta saran untuk penelitian selanjutnya.

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB II

LANDASAN TEORI

2.1. *Natural Language Prosessing* (NLP)

Natural Language Prosessing (NLP) adalah salah satu bidang ilmu Kecerdasan Buatan atau disebut juga dengan *Artificial Intelligence* yang mempelajari komunikasi antara manusia dengan komputer melalui bahasa alami. NLP tidak bertujuan untuk mentransformasikan bahasa yang diterima dalam bentuk suara menjadi data digital dan atau sebaliknya pula, melainkan bertujuan untuk memahami arti dari ucapan yang diberikan dalam bahasa alami dan memberikan respon yang sesuai, misalnya dengan melakukan suatu aksi tertentu atau menampilkan data tertentu (James Suciadi Industri, 2004).

Natural Language Processing (NLP) merupakan salah satu cabang ilmu AI yang berfokus pada pengolahan bahasa natural. Bahasa natural adalah bahasa yang secara umum digunakan oleh manusia dalam berkomunikasi satu sama lain. Bahasa yang diterima oleh komputer butuh untuk diproses dan dipahami terlebih dahulu supaya maksud dari *user* bisa dipahami dengan baik oleh komputer.

Ada berbagai terapan aplikasi dari NLP. Diantaranya adalah *Chatbot* (aplikasi yang membuat *user* bisa seolah-olah melakukan komunikasi dengan komputer), *Stemming* atau *Lemmatization* (pemotongan kata dalam bahasa tertentu menjadi bentuk dasar pengenalan fungsi setiap kata dalam kalimat), *Summarization* (ringkasan dari bacaan), *Translation Tools* (menterjemahkan bahasa) dan aplikasi-aplikasi lain yang memungkinkan komputer mampu memahami instruksi bahasa yang diinputkan oleh *user*.

Pustejovsky dan Stubbs (2012) menjelaskan bahwa ada beberapa area utama penelitian pada *field* NLP, diantaranya:

1. *Question Answering Systems* (QAS). Kemampuan komputer untuk menjawab pertanyaan yang diberikan oleh *user*. Daripada memasukkan *keyword* ke dalam *browser* pencarian, dengan QAS, *user* bisa langsung bertanya dalam bahasa natural yang digunakannya, baik itu Inggris, Mandarin, ataupun Indonesia.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. *Summarization*. Pembuatan ringkasan dari sekumpulan konten dokumen atau *email*. Dengan menggunakan aplikasi ini, *user* bisa dibantu untuk mengkonversikan dokumen teks yang besar ke dalam bentuk *slide* presentasi.
3. *Machine Translation*. Produk yang dihasilkan adalah aplikasi yang dapat memahami bahasa manusia dan menterjemahkannya ke dalam bahasa lain. Termasuk di dalamnya adalah *Google Translate* yang apabila dicermati semakin membaik dalam penterjemahan bahasa. Contoh lain lagi adalah *BabelFish* yang menterjemahkan bahasa pada *real time*.
4. *Speech Recognition*. *Field* ini merupakan cabang ilmu NLP yang cukup sulit. Proses pembangunan model untuk digunakan telpon/komputer dalam mengenali bahasa yang diucapkan sudah banyak dikerjakan. Bahasa yang sering digunakan adalah berupa pertanyaan dan perintah.
5. *Document classification*. Sedangkan aplikasi ini adalah merupakan area penelitian NLP Yang paling sukses. Pekerjaan yang dilakukan aplikasi ini adalah menentukan dimana tempat terbaik dokumen yang baru diinputkan ke dalam sistem. Hal ini sangat berguna pada aplikasi *spam filtering*, *news article classification*, dan *movie review*.

Perkembangan NLP menghasilkan kemungkinan dari *interface* bahasa natural menjadi *knowledge base* dan penterjemahan bahasa natural. Poole dan Mackworth (2010) menjelaskan bahwa ada 3 (tiga) aspek utama pada teori pemahaman mengenai *natural language*:

1. *Syntax*: menjelaskan bentuk dari bahasa. *Syntax* biasa dispesifikasikan oleh sebuah *grammar*. *Natural language* jauh lebih daripada formal *language* yang digunakan untuk logika kecerdasan buatan dan program komputer
2. *Semantics*: menjelaskan arti dari kalimat dalam satu bahasa. Meskipun teori *semantics* secara umum sudah ada, ketika membangun sistem *natural language understanding* untuk aplikasi tertentu, akan digunakan representasi yang paling sederhana.
3. *Pragmatics*: menjelaskan bagaimana pernyataan yang ada berhubungan dengan dunia. Untuk memahami bahasa, agen harus mempertimbangan lebih dari hanya

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

sekedar kalimat. Agen harus melihat lebih ke dalam konteks kalimat, keadaan dunia, tujuan dari *speaker* dan *listener*, konvensi khusus, dan sejenisnya.

Contoh kalimat di bawah ini akan membantu untuk memahami perbedaan diantara ketiga aspek tersebut di atas. Kalimat-kalimat ini adalah kalimat yang mungkin muncul pada bagian awal dari sebuah buku Artificial Intelligence (AI):

1. *This book is about Artificial Intelligence*
2. *The green frogs sleep soundly*
3. *Colorless green ideas sleep furiously*
4. *Furiously sleep ideas green colorless*

Kalimat pertama akan tepat jika diletakkan pada awal sebuah buku, karena tepat secara sintaks, semantik, dan pragmatik. Kalimat kedua tepat secara sintaks dan semantic, namun kalimat tersebut akan menjadi aneh apabila diletakkan pada awal sebuah buku AI, sehingga kalimat ini tidak tepat secara pragmatik. Kalimat ketiga tepat secara sintaks, tetapi tidak secara semantik. Sedangkan pada kalimat keempat, tidak tepat secara sintaks, semantik, dan pragmatik.

Selain daripada ketiga istilah tersebut ada beberapa istilah yang terkait dengan NLP, yaitu:

1. Morfologi merupakan pengetahuan tentang kata dan bentuknya sehingga bisa dibedakan antara yang satu dengan yang lainnya. Bisa juga didefinisikan asal usul sebuah kata itu bisa terjadi.
Contoh : membangunkan → bangun (kata dasar), mem- (*prefix*), -kan (*suffix*)
2. *Fonetik* merupakan segala hal yang berhubungan dengan suara yang menghasilkan kata yang dapat dikenali. *Fonetik* digunakan dalam pengembangan NLP khususnya bidang *speech based system*.

2.2. Bahasa Minang

Sebagai sarana pikir, ekspresi, dan komunikasi, bahasa senantiasa berkembang sejalan dengan perkembangan peradaban masyarakat pendukungnya. Perkembangan bahasa itu akan tampak dari penambahan kosakata dari waktu ke waktu. Pada permulaan abad ke-15 tercatat 500 lema bahasa Melayu dalam Daftar

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Kata Cina-Melayu (dokumen sejarah perkamusan Indonesia). Pada pertengahan abad ke-20, tercatat sekitar 23.000 lema dalam *Kamus Umum Bahasa Indonesia Poerwadarminta* (1953). Pada penerbitan tahun 1976 (dalam kurun waktu 23 tahun) lema dalam kamus itu menjadi 24.000. Dua belas tahun kemudian (tahun 1988) terbit *Kamus Besar Bahasa Indonesia Edisi Pertama* yang memuat 62.000 lema. Perkembangan yang amat pesat itu terus berjalan, terutama pada perempat abad ke-20 terakhir dan awal abad ke-21 ini. Pada awal abad ini *Kamus Besar Bahasa Indonesia Pusat Bahasa Edisi IV* (2008) memuat hampir 91.000 lema. Di samping kata umum, dalam bahasa Indonesia telah dikembangkan istilah berbagai bidang ilmu dan teknologi serta seni yang kini telah dihasilkan sekitar 405.000 istilah. Kini terus dikembangkan kosakata dan istilah, baik yang bersumber dari bahasa asing maupun yang bersumber dari bahasa daerah di seluruh nusantara ini, termasuk dari bahasa Minangkabau. Semua itu merupakan kekayaan yang memperlihatkan kemajuan peradaban bangsa Indonesia (Burhanuddin et al., 2012).

Istilah Minangkabau mengacu pada batasan budaya bukan merupakan batasan geografis suatu wilayah. Dari segi budaya, Minangkabau terletak di bagian tengah pulau Sumatra dan sebagian semenanjung Malaysia (Naim, 1984), yaitu yang meliputi

- 1) Provinsi Sumatra Barat, minus Mentawai, 2) Riau, 3) bagian utara Provinsi Jambi, 4) bagian utara Bengkulu, 5) pantai barat Nanggroe Aceh Darussalam dan Sumatra Utara, serta 6) Negeri Sembilan, Malaysia (Burhanuddin et al., 2012).

Sumatra Barat yang merupakan daerah Asal Minangkabau, terbagi atas dua bagian yaitu *darek* dan *rantau*. *Darek* adalah daerah tua Minangkabau, yang terdiri atas tiga *luhak*, yaitu Luhak Tanahdatar, Luhak Agam, dan Luhak Limo Puluah Koto. Ketiga daerah tersebut terletak di sekitar Gunung Marapi, Singgalang, dan Sago. Daerah *rantau* meliputi pesisir pantai Sumatra, di antaranya Padang, Pariaman, Pesisir Selatan, Kerinci, Natal, dan Indera Giri (Burhanuddin et al., 2012).

Bahasa Minangkabau merupakan sarana komunikasi bagi masyarakat minang. Bahasa ini memiliki berbagai sebutan, yaitu bahasa Minang atau bahasa Padang (Grimes, dalam Nadra, 2006) dan bahasa Melayu Minangkabau (Muhajir, dalam Nadra, 2006). Jumlah penutur bahasa Minangkabau menduduki peringkat kelima dari

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

sepuluh bahasa daerah terbesar yang ada di Indonesia. Di Sumatra, bahasa Minangkabau adalah bahasa kedua terbesar setelah bahasa Melayu (Burhanuddin et al., 2012).



Gambar 2. 1 Cover Buku Kamus Bahasa Minangkabau-Indonesia

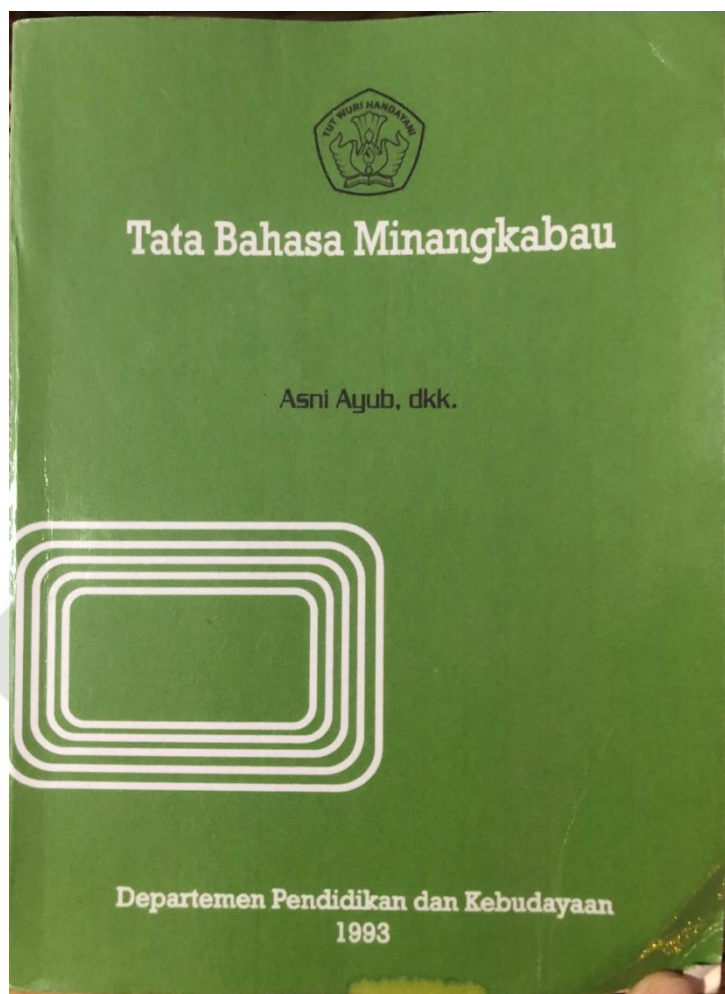
2.3. Pola Bahasa Minangkabau

Secara fonologi, Moussay Gerard, (1998) menyatakan bahwa bahasa Minangkabau mencakupi 19 konsonan dan 5 vokal. Bahasa Minangkabau pada dasarnya tidak mengenal abjad f dan h. Kedua huruf tersebut tergabung dalam huruf p dan a. Disamping itu, bahasa ini juga tidak mengenal penggunaan huruf q dan z yang dalam sehari-hari memakai huruf awal k dan j dan tidak pernah mengenal huruf x. Akan tetapi karena pengaruh agama Islam dan hubungan antar suku bukan saja di

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Indonesia, masyarakat Minangkabau mengenal penggunaan huruf q dan z seperti terlihat dalam ucapan: qadar, qurban, zakat, ziarah dan lain sebagainya.



Gambar 2. 2 Cover Buku Tata Bahasa Minangkabau

Huruf e yang hampir selalu terpakai dalam Bahasa Minangkabau adalah e pepet seperti dalam kosa kata: ameh, Aceh, kameh, mameh, bareh, areh, lapeh, leleh dan lain sebagainya. Berikut merupakan pola bahasa Minangkabau atau tata aturan perubahan kata yaitu:

1. Untuk setiap kata tanya hilangkan satu suku kata terakhir.

Contoh:

- a. “Siapa” = “sia (pa)”
- b. “Apa” = “a (pa)”
- c. “Berapa” = “bara (pa)”

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

d. “Mengapa” = “manga (pa)”

e. “Dimana” = “dima (na)”

Dalam bahasa Minangkabau ada konsep dasar yang harus selalu dipegang, yaitu: tidak terdapat awalan menggunakan vokal (e) sehingga untuk setiap awalan menggunakan huruf vokal “e” diganti menjadi (a), contohnya (berapa) tidak menjadi (bera) tetapi menjadi (bara).

Memang tidak semua kata tanya, ada beberapa kata tanya yang tidak sesuai pola seperti:

a. “Bagaimana” = “baa”

b. “Kapan” = “bilo”

2. Setiap bentuk kata yang ujung suku katanya “as” berubah menjadi “eh”

Contoh:

a. “Lepas” = “lapeh”

b. “Beras” = “bareh”

c. “Lemas” = “lameh”

d. “Kuras” = “kureh”

e. “Atas” = “ateh”

f. “Bekas” = “bakeh”

g. “Luas” = “lueh”

3. Setiap bentuk kata yang ujung suku katanya “at” berubah menjadi “ek”

Contoh:

a. “Berat” = “barek”

b. “Kuat” = “kuek”

c. “Lambat” = “lambek”

d. “Lebat” = “labek”

e. “Buat” = “buek”

f. “Tempat” = “tampek”

g. “Bulat” = “bulek”

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

4. Setiap bentuk kata yang ujung suku katanya "ap" berubah menjadi "ok"

Contoh:

- a. "Atap" = "atok"
- b. "Hisap" = "hisok"
- c. "Suap" = "suok"
- d. "Uap" = "uok"

5. Setiap kata yang ujung suku katanya "ing" ditambahkan huruf a, menjadi "iang"

Contoh:

- a. "Kucing" = "kuciang"
- b. "Maling" = "maliang"
- c. "Runcing" = "runciang"
- d. "Kambing" = "kambiang"

6. Setiap kata yang ujung suku katanya "ung" ditambahkan huruf a, menjadi "uang"

Contoh:

- a. "Untung" = "untuang"
- b. "Burung" = "buruang"
- c. "Hitung" = "hituang"
- d. "Kampung" = "kampuang"
- e. "Sarung" = "saruang"

7. Setiap kata yang ujung suku katanya "uh" ditambahkan huruf a, menjadi "uah"

Contoh:

- a. "Rusuh" = "rusuah"
- b. "Kumuh" = "kumuah"
- c. "Rubuh" = "rubuah"
- d. "Patuh" = "patuah"
- e. "Tumbuh" = "tumbuah"

8. Setiap kata yang ujung suku katanya "us" berubah menjadi "uih"

Contoh:

- a. "Putus" = "putuih"
- b. "Pupus" = "pupuiah"

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- c. “Hapus” = “hapuih”
- d. “Hembus” = “hambuih”
- e. “Kurus” = “kuruih”
- f. “Tembus” = “tambuih”

9. Setiap kata yang ujung suku katanya “it” berubah menjadi “ik”

Contoh:

- a. “Sakit” = “sakik”
- b. “Pelit” = “pilik”
- c. “Tumit” = “tumik”
- d. “Rakit” = “rakik”
- e. “Rumit” = “rumik”
- f. “Himpit” = “himpik”

10. Setiap kata yang ujung suku katanya “is” berubah menjadi “ih”

Contoh:

- a. “Manis” = “manih”
- b. “Tulis” = “tulih”
- c. “Baris” = “barih”
- d. “Habis” = “habih”
- e. “Gadis” = “gadiah”

11. Setiap kata yang ujung suku katanya “ut” berubah menjadi “uik”

Contoh:

- a. “Lutut” = “lutuik”
- b. “Lumut” = “lumuik”
- c. “Rambut” = “rambuik”
- d. “Kusut” = “kusuik”
- e. “Takut” = “takuik”
- f. “Urut” = “uruik”
- g. “Belut” = “baluik”

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

12. Setiap kata yang berakhiran “a” berubah menjadi “o”

Contoh:

- | | |
|-------------|------------|
| a. “Bunga” | = “bungo” |
| b. “Cinta” | = “cinto” |
| c. “Suka” | = “suko” |
| d. “Gula” | = “gulo” |
| e. “Mata” | = “mato” |
| f. “Kepala” | = “kapalo” |
| g. “Dada” | = “dado” |
| h. “Ada” | = “ado” |
| i. “Luka” | = “luko” |

13. Setiap kata yang berakhiran “ur” berubah menjadi “ua”

Contoh:

- | | |
|------------|-----------|
| a. “Kubur” | = “kubua” |
| b. “Cukur” | = “cukua” |
| c. “Ukur” | = “ukua” |
| d. “Kasur” | = “kasua” |

14. Jika ada kata terakhir dengan huruf “r” maka huruf terakhir dibuang.

Contoh:

- | | |
|------------|----------|
| a. “Bakar” | = “baka” |
| b. “Kasar” | = “kasa” |
| c. “Tukar” | = “tuka” |
| d. “Datar” | = “data” |
| e. “Pagar” | = “paga” |
| f. “Putar” | = “puta” |

15. Partikel “do” dan “mah”

Partikel “do” dan “mah” akan sering dijumpai dalam akhiran kalimat bahasa Minangkabau. Partikel tersebut tidak memiliki arti akan tetapi sangat penting dalam penegasan kalimat. Partikel “do” biasanya dalam penegasan kalimat *negatif*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

(bermakna tidak), sedangkan partikel ‘mah’ biasanya dalam penegasan kalimat sanggahan/tuduhan/perkiraan kuat yang sudah terjadi (*present perfect tense*).

Contoh :

- a. “ndak ado jo den do” = “nggak ada sama aku (berarti penegasan bahwa barang tersebut benar-benar tidak ada di saya)”.
- b. “alun baraja lai do” = “belum belajar nih (berarti penegasan bahwa orang tersebut memang benar-benar belum belajar)”.
- c. “ndak tahu do” = “nggak tahu (berarti benar-benar tidak tahu)”.
- d. “bukunyo jo waang mah” = “kan buku dia sama kamu (lebih penegasan menuduh buku tersebut ada di kamu)”.
- e. “lah den agiah ang pitih mah” = “kan udah aku kasih kamu duit (penegasan bahwa uang itu sudah diberikan)”.

2.4. Morfologi Bahasa Minang

Menurut definisi (Samsuri, 1978:190) proses morfologis adalah cara pembentukan kata-kata dengan menghubungkan morfem yang satu dengan morfem lainnya. Karena kita menghubungkan-morfem-morfem menjadi kata, maka morfem merupakan bentuk morfologis terkecil, sedangkan yang terbesar adalah kata. Didalam BMU ada dua proses morfologis dalam pembentukan kata, yaitu pengimbuhan (afiksasi) dan perulangan (reduplikasi).

2.4.1. Imbuhan dan Distribusinya

BMU (Bahasa Minang Umum) mempunyai imbuhan dalam bentuk awalan, akhiran dan gabungan awalan akhiran. Awalan banyak sekali jumlahnya dan amat produktif dan lebih rumit, sedangkan akhiran sedikit sekali jumlahnya tetapi mempunyai fungsi yang besar dan amat produktif. Selain dari bisa berdiri sendiri dalam pembentukan kata, awalan dan akhiran dapat bergabung dan berganda, sehingga terjadi pula proses imbuhan baru dengan distribusi dan arti yang baru.

a. Awalan (Prefiks)

Dalam BMU terdapat bentuk-bentuk awalan sebagai berikut: / ba-1 /, / ba-2 /, /ba-3/, / maN- /, / paN- /, /pa- /, /ta-1/, /ta-2/, /di /, / sa- /, / ka- /.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

1. Awalan / ba-1 /
Awalan / ba-1 / ini dapat bergabung dengan nomina dan gabungan ini membentuk verba intransitif. Dengan penggabungan ini, prefiks ini memiliki beberapa arti, yaitu: mempunyai, memakai, mengendarai, menghasilkan, memanggil/menyebut dan bersifat/berkemauan.
Contoh: / bapaga / ‘berpagar’ / babaju / ‘berbaju’
2. Awalan / ba-2 /
Awalan / ba-2 / ini tidak produktif dan hanya terdapat pada kata baraja ‘belajar’ yang merupakan verba transitif.
Contoh: / adiak sedang baraja baretong / ‘adik sedang belajar berhitung’
3. Awalan / ba-3 /
Awalan / ba-3 / ini hanya bergabung dengan verba transitif dan mengandung makna pasif.
Contoh: / Lah batulih surek itu? / ‘apakah surat itu telah ditulis?’
4. Awalan / maN- /
Awalan / maN- (mam-, man-, mang-, many- dan ma-) / adalah pembentuk Kata Kerja, baik transitif maupun intransitif. Dalam pembentukan kata kerja ini digunakan Kata Kerja, Kata Benda atau Kata Sifat sebagai kata dasar.
Contoh: / mananak nasi / ‘memasak nasi’
/ marauang / ‘meraung’ / marakik / ‘membuat rakit’
/ maijaw / ‘menjadi hijau’ / maitam / ‘menjadi hitam’
5. Awalan / paN- /
Awalan / paN- (pa-, pam-, pan-, pang- dan pany-) / dapat dihubungkan dengan Kata Kerja, Kata Benda atau Kata Sifat sebagai kata dasar sehingga terbentuk kata-kata baru dengan pengertian baru pula.
Contoh: / paota / ‘suka ngobrol’ / palapeh / ‘sering lepas’
/ pancalieh / ‘suka malas’ / pangareh / ‘suka bersikeras’

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

6. Awalan / pa- /
Awalan / pa- / dapat dihubungkan dengan Kata Sifat, Kata Benda atau Kata Bilangan yang menjadikan Kata Kerja.
Contoh: / pabudak / ‘menjadikan budak’
 7. Awalan / ta-1/
Awalan / ta-1 / yang mengawali Kata Sifat, Kata Kerja atau Kata Benda membentuk kata jadian dengan pengantin baru.
Contoh: / tacampak / ‘terbuang’ / tamakan / ‘termakan’
 8. Awalan /ta-2/
Awalan /ta-2/ ini bergabung dengan adjektiva atau verba intransitive yang jumlahnya sangat terbatas.
Contoh: /tapaik/ ‘rasa pahit’
 9. Awalan / di- /
Awalan /di-/ ini melekat pada verba transitif saja.
Contoh: /dibaco/ ‘baca’ /diambiak/ ‘ambil’
 10. Awalan / sa- /
Awalan / sa- / dapat digabungkan dengan Kata Sifat, Kata Benda atau Kata Bilangan.
Contoh: / satinggi rumah / ‘sama tinggi dengan rumah’
 11. Awalan / ka- /
Awalan / ka- / dapat dikaitkan dengan Kata Bilangan dengan pengertian urutan’ dan dapat diartikan untuk menunjukkan arah.
Contoh: / ambo anak nan ka ampek / ‘saya adalah anak yang ke empat’
/ kalua jan masuak juo / “keluar jangan masuk juga”
- b. Akhiran (Sufiks)**
- Akhiran dalam BMU ada tiga yakni /-an/, /-K.an/, /-i/.
1. Akhiran /-an/
Akhiran /-an/ dapat ditambahkan pada Kata Kerja, Kata Benda dan Kata Sifat yang membuat kata tersebut menjadi Kata benda derivatif.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2. Awalan Akhiran /ba-.....-an/

Imbuhan awalan akhiran /ba-.....-an/ merupakan verba atau numeralia.

Contoh: /datang/ 'datang' /badatangan/ 'berdatangan'

/lari/ 'lari' /balarian/ 'berlarian'

3. Awalan Akhiran /ba-.....-i/

Imbuhan awalan akhiran /ba-.....-i/ hanya bergabung dengan verba transitif.

Konfiks ini juga merupakan konfiks yang infleksional karena tidak mengubah kelas kata yang dilekatinya.

Contoh: /bae/ 'lempar' /babaei/ 'dilempari'

/tokok/ 'pukul' /batokoki/ 'dipukuli'

4. Awalan Akhiran /ka-.....-an/

Imbuhan awalan akhiran /ka-.....-an/ ini bergabung dengan verba atau adjective. Kata bentukan yang dihasilkan adalah nomina, karena itu konfiks ini menunjukkan hasil dari kata yang dilekatinya.

Contoh: /duduak/ 'duduk' /kaduduakan/ 'kedudukan'

/datang/ 'datang' /kاداتangan/ 'kedatangan'

2.4.2. Morfem

Sebuah morfem dalam BMU bisa terdiri dari satu silabe, dua silabe, tiga silabe, atau empat silabe, tetapi tidak ada morfem yang bersuku lebih dari empat. Sedangkan sebuah suku kata bisa terdiri dari sebuah vokal, atau sebuah vokal di dahului oleh sebuah konsonan, atau sebuah vokal diikuti oleh sebuah konsonan, atau sebuah vokal di dahului dan diikuti oleh sebuah konsonan.

Contoh-contoh:

a. Morfem bersuku satu:

/nan / 'yang' / ei / '.hei', / jam / 'jam', / -an / (sebuah akhiran), / ba- / (sebuah awalan).

b. Morfem bersuku dua:

/ i - no / 'ia', / su-re? / 'surat', /gun-tiang-, / 'gunting', /basi-/ (sebuah awalan).

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

- c. Morfem bersuku tiga:

/ ka-pa-tan / 'kemarin', / te-la-pa? / 'telapak', / ga-re-nah /, "prilaku', seng.gu-lung-
, / 'senggulung'.
- d. Morfem bersuku empat:

/ma-ra-pu-lay / 'mempelai', / ka-la-la-wa / 'kelelawar', / ka-titi-ti-ran / 'burung
ketitiran', / ka-U-ma-yia / 'sebangsa ulat'

2.4.3. Klasifikasi Kata

Secara morfemis, kata-kata BMU dibagi atas dua macam, yaitu kata yang terdiri dari satu morfem (mono-morfem) dan kata yang terdiri dari lebih dari satu morfem (poli-morfem). sedangkan secara sintaksis kata dibagi atas kelas-kelas benda, kerja, sifat, keterangan, bilangan, depan, sandang, sambung, dan seru.

a. Kata Mono-Morfem

Sebuah morfem yang dapat berdiri sendiri tanpa tambahan morfem lain disebut kata mono-morfem, seperti:

* ado / 'ada', / maa / 'mana', /wa?ang / 'engkau laki-laki', / bendi / 'dokar', / saki? / 'sakit', elo? / 'baik', / karajo / 'kerja', /de? / 'karena.

b. Kata Polimorfem

Dalam BMU kata-kata berikut adalah kata polimorfem:

/gulo/ 'gula' menjadi /gulo-gulo/ 'gula-gula' /kudo/ 'kuda' menjadi /kudo-kudo/ 'kuda-kuda' /gali/ 'geli' menjadi /gali-gali/ 'geli' /lari/ 'lari' menjadi /lari-lari/ 'lari-lari'

Kata poli-morfem juga dapat terbentuk dari dua monomorfem yang berbeda, umpamanya :

/jantuang/ - ati / 'kekasih' /urang-mudo/ 'pemuda'
/itam-putih/ 'menguasai' /kacang-goreng/ 'kacang tanah'

Selain dari dua tipe di atas, kata poli-morfem bisa juga dibentuk dari kata mono-morfem atau poli-morfem dengan imbuhan, misalnya:

/mang-galeh / 'berjualan
/ ba-siram / 'bersiram atau disiram'.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

2.4.4. Proses Morfologis

Menurut definisi (Samsuri, 1978: 190) proses morfologis ialah cara pembentukan kata-kata dengan menghubungkan morfem yang satu dengan morfem yang lain. Karena kita menghubungkan-menghubungkan morfem-morfem menjadi kata, maka morfem merupakan bentuk morfologis terkecil, sedangkan yang terbesar adalah kata. Di dalam BMU ada dua proses morfologis dalam pembentukan kata, yaitu pengimbuhan (afiksasi) dan perulangan (reduplikasi).

a. Imbuhan dan Distribusinya

BMU mempunyai imbuhan dalam bentuk awalan, akhiran dan gabungan awalan akhiran. Awalan banyak sekali jumlahnya dan amat produktif dan lebih rumit; sedangkan akhiran sedikit sekali jumlahnya tetapi mempunyai fungsi yang besar dan amat produktif.

(1) Awalan (Prefiks)

Dalam BMU terdapat bentuk-bentuk awalan sebagai berikut: /ba-1 /, /ba-2 /, /ba-3/, /maN- /, /paN-/, /pa- /, /ta-1/, /ta-2/ /di /, /ka- /.

(2) Akhiran (Sufiks)

Akhiran dalam BMU ada empat yakni /-an/, /-K.an/, /-i/.

(3) Awalan Akhiran (Konfiks)

Dalam BMU terdapat bentuk-bentuk gabungan imbuhan Awalan akhiran yaitu sebagai berikut: /PaN-.....-an/, /Ba-.....-an/, /Ba-.....-i/ dan /Ka-.....-an/.

b. Reduplikasi

Reduplikasi ialah proses pembentukan kata dengan perulangan kata dasar. Reduplikasi bisa dialami oleh kata dasar baik secara keseluruhan atau sebahagian dengan atau tanpa perobahan sebuah fonem atau lebih dari 'bagian kata yang diulang. Kata berulang bisa pula mendapat imbuhan. Jadi reduplikasi dalam BMU dapat dibagi atas tiga bagian: reduplikasi sempurna, reduplikasi tidak sempurna, dan reduplikasi berimbuhan.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

(1) Reduplikasi Sempurna

Reduplikasi sempurna dalam BMU terdapat dengan KB, KK, KS, dan KBil. Reduplikasi pada KK dan KS berarti intensitas pekerjaan dan sifat, dengan bilangan berarti kelompok dan menjadikannya KKet, dan dengan KB berarti banyak atau menyerupai.

(2) Reduplikasi tidak Sempurna

Terbagi menjadi 3 bagian yaitu dengan perubahan vokal pada kata kedua, dengan perubahan atau penambahan konsonan pada kata kedua, dan yang tidak beraturan.

(3) Reduplikasi dengan imbuhan

Proses reduplikasi baik dengan imbuhan maupun tidak dapat disimpulkan dalam tabel di bawah ini:

<maN>	KK	-	<R>	-	<lah>
-	KK	<maN>	<R>	-	<lah>
<ba ₁ >	KK	-	<R>	<kan>	<lah>
<pa>	KK	-	<R>	<kan>	<lah>
<ba ₁ > <pa/>	KS	-	<R>	<kan>	<lah>
<ta> <pa>	KS	-	<R>	-	<lah>
<ba ₂ >	KB	-	<R>	-	<lah>
<ba ₂ >	KBil	<ba ₂ >	<R>	-	<lah>

Tabel 2. 1 Reduplikasi dan Imbuhan

2.5. Stemming

Stemming merupakan suatu proses untuk menemukan kata dasar dari sebuah kata. *Stemming* merupakan proses yang memetakan bentuk varian kata menjadi kata dasarnya (Tala, 2004). Untuk *stemming* dalam bahasa Indonesia, terdapat dua jenis metode *stemming* yang sudah ada, yaitu algoritma *stemming* yang berbasis kamus (*dictionary based*) dan algoritma *stemming* yang berbasis non-kamus (*purely rule*

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

based). *Stemming* merupakan proses yang digunakan untuk menemukan kata dasar dari sebuah kata dengan menghilangkan semua imbuhan (*affixes*) baik yang merupakan awalan (*prefixes*), akhiran (*suffixes*), dan kombinasi dari awalan serta akhiran (*confixes*) pada kata turunan (Manase Sahat H Simarangkir, 2017).

2.6. Penelitian Terkait

Pada Tabel 2.1 berikut ini menjelaskan tentang penelitian terkait yang telah dilakukan sebelumnya.

Tabel 2. 2 Penelitian Terkait

No	Nama penulis	Judul	Thn	Metode	Hasil
1	Ledy Agusta	“Perbandingan Algoritma Stemming Porter dengan Algoritma Nazief & Adriani untuk Stemming Dokumen Teks Bahasa Indonesia”	2009	Algoritma Stemming Porter Dengan Algoritma Nazief & Adriani	“Proses stemming dokumen teks berBahasa Indonesia menggunakan Algoritma Porter membutuhkan waktu yang lebih singkat dibandingkan dengan stemming menggunakan Algoritma Nazief & Adriani. Proses stemming dokumen teks berBahasa Indonesia menggunakan Algoritma Porter memiliki prosentase keakuratan (presisi) lebih kecil dibandingkan dengan stemming menggunakan Algoritma Nazief & Adriani. Pada proses stemming menggunakan Algoritma Nazief & Adriani, kamus yang

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Nama penulis	Judul	Thn	Metode	Hasil
					digunakan sangat mempengaruhi hasil stemming. Semakin lengkap kamus yang digunakan maka semakin akurat pula hasil stemming”
2	Andhy Purwoko	“Model Stemming Berbasis Kamus untuk Dokumen Berbahasa Sunda”	2011	Algoritma <i>stemming</i>	“Algoritma <i>stemming</i> yang dirancang mampu memberikan hasil kata yang diharapkan sebesar 91.38% dan mengurangi jumlah kata pada indeks sebesar 49.31%”
3	Ananda Wulandari, Kemas Rahmat S.W, Ade Romadhony	“Pattern-Based Stemmer Analysis And Implementati on On Arabic Text”	2011	<i>Pattern-Based Stemmer</i>	“Pattern-based stemmer adalah algoritma yang cukup efisien untuk menentukan akar kata dari teks-teks Arab”
4	Lasmedi Afuan	“Stemming Dokumen Teks Bahasa Indonesia Menggunakan Algoritma Porter”	2013	Algoritma Porter	“Presisi pada Proses <i>stemming</i> masih belum mencapai hasil yang maksimal, hal ini bukan karena aplikasi yang tidak benar, akan tetapi kamus kata dasar yang masih belum terlalu lengkap. Aplikasi Stemming dokumen bahasa indonesia ini, sementara hanya bisa

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Nama penulis	Judul	Thn	Metode	Hasil
					membaca dokumen dengan ekstensi (.txt), sehingga aplikasi ini masih harus disempurnakan agar bisa membaca berbagai format dokumen”
5	Damar Aji Asmara, Dewi Khairani, Siti Umami Masruroh	“Penerapan Algoritma Paice atau Husk untuk Stemming Pada Kamus Bahasa Inggris Ke Bahasa Indonesia”	2013	Algoritma Paice/Husk Dan Porter	“Tingkat akurasi algoritma Paice/Husk sebesar 98.3% sedangkan algoritma Porter sebesar 55.6%. Dengan demikian, algoritma Paice/Husk lebih cocok untuk penelusuran kata dasar pada kamus bahasa Inggris”
6	Noverdy Anggara, Ade Romadhony, Mahmud Dwi Suliyono	“Implementasi Modifikasi Algoritma Enhanced Confix Stripping Stemmer Pada Teks Bahasa Indonesia”	2013	Algoritma Enhanced Confix Stripping Stemmer	“Algoritma Enhanced Confix Stripping Stemmer dapat digunakan untuk proses Stemming dokumen text berbahasa Indonesia. Tetapi beberapa kata tidak mampu distemming dikarenakan kata gabung, bahasa asing, salah tulis dan kata tidak tertera dalam kamus. Modifikasi algoritma ECS Stemmer memiliki akurasi yang lebih tinggi

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Nama penulis	Judul	Thn	Metode	Hasil
					dibandingkan algoritma ECS murni dengan perbedaan akurasi sebesar 0,1% sampai dengan 5,0%”
7	Rakhmad Maulidi	“ <i>Stemmer</i> untuk Bahasa Madura dengan Modifikasi Metode <i>Enhanced Confix Stripping Stemmer</i> ”	2016	<i>Enhanced Confix Stripping Stemmer</i>	“Metode <i>Enhanced Confix Stripping Stemmer</i> memungkinkan digunakan untuk stemmer bahasa Madura dengan melakukan penyesuaian pada <i>rule base</i> -nya sesuai dengan morfologi bahasa Madura”
8	Julianto Wibowo	“Aplikasi Penentu Kata Dasar dari Kata Berimbuhan Pada Kalimat Bahasa Indonesia dengan Algoritma Stemming”	2016	Algoritma Stemming Nazief dan Adriani	“Cara kerja pada algoritma <i>stemming</i> Nazief dan Adriani berupa pengambilan langkah pemotongan pertama pada semua akhiran terlebih dahulu dan selanjutnya dilakukan pemotongan awalan serta pengecekan database kata dasar pada awal dan akhir proses <i>stemming</i> ”
9	Mar’atus Madia	“Stemming Bahasa Jawa untuk Mencari	2016	Algoritma <i>stemming</i> , menggunakan aturan	“Akurasi yang didapatkan sebesar 92,881% dari 16.745 kata bahasa Jawa”

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Nama penulis	Judul	Thn	Metode	Hasil
		Akar Kata dalam Bahasa Jawa dengan Aturan Analisis Kontrasif Afiksasi Verba”		analisis kontrasif afiksasi verba	
10	Dian Novitasari	“Perbandingan Algoritma Stemming Porter dengan Arifin Setiono untuk Menentukan Tingkat Ketepatan Kata Dasar”	2016	Algoritma Stemming Porter dengan Arifin Setiono	“Hasil <i>stemmer Exact Match</i> sebesar 90% untuk <i>stemmer Porter</i> dan 95% untuk <i>stemmer Arifin Setiono</i> , <i>Unchange</i> sebesar 2,5% untuk <i>stemmer Porter</i> dan 2,5% untuk <i>stemmer Arifin Setiono</i> , <i>Spelling Exception</i> sebesar 7,5% untuk <i>stemmer Porter</i> dan 0% untuk <i>stemmer Arifin Setiono</i> , dan <i>Overstemming</i> sebesar 0% untuk <i>stemmer Porter</i> dan 2,5% untuk <i>stemmer Arifin Setiono</i> ”
11	Gusti Ngurah Mega Nata, Putu Pande Yudiastra	“ <i>Stemming</i> teks <i>sor-singih</i> Bahasa Bali”	2017	Algoritma <i>stemmer</i>	“Dari hasil pengujian 85% kata <i>distemming</i> dengan benar”
12	Dwi Wahyudi, Teguh	“Implementasi dan	2017	Algoritma Nazief dan	“Dari hasil tersebut tampak, bahwa algoritma Nazief &

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Nama penulis	Judul	Thn	Metode	Hasil
	Susyanto, Didik Nugroho	Analisis Algoritma Stemming Nazief & Adriani Dan Porter Pada Dokumen Berbahasa Indonesia”		Adriani dengan Algoritma Porter	Adriani (Akurasi 95,26% , Waktu proses 22.1668348312) memberikan hasil yang lebih baik untuk <i>stemming</i> pada dokumen berbahasa Indonesia. Namun untuk efisiensi waktu proses, algoritma Porter (Akurasi 79,13% , Waktu 12.3822753429) lebih baik dibanding algoritma Nazief & Adriani”
13	Manase Sahat H Simarangkir	“Studi Perbandinga n Algoritma – Algoritma Stemming untuk Dokumen Teks Bahasa Indonesia”	2017	Algoritma Nazief dan Adriani, Algoritma Tala	“Algoritma terbaik dalam proses <i>stemming</i> yaitu proses algoritma Nazief dan Adriani untuk algoritma yang menggunakan kamus. Sedangkan untuk algoritma yang memakai aturan imbuhan ditemukan algoritma terbaik dalam proses <i>stemming</i> pada algoritma Tala”
14	Adhi Prasidhatama	“Analisis Perbandinga n Algoritma Nazief dan Adriani dengan Algoritma	2017	Algoritma Nazief dan Adriani, Algoritma Idris	“Algoritma <i>stemming</i> Idris memiliki waktu pemrosesan lebih cepat dibandingkan dengan algoritma <i>stemming</i> Nazief dan Adriani. Algoritma <i>stemming</i> Nazief dan

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

No	Nama penulis	Judul	Thn	Metode	Hasil
		Idris untuk Pencarian Kata Dasar Pada Teks Bahasa Indonesia Berbasis <i>Website</i> ”			Adriani memiliki akurasi hasil yang lebih tinggi dibandingkan dengan algoritma <i>stemming</i> Idris”



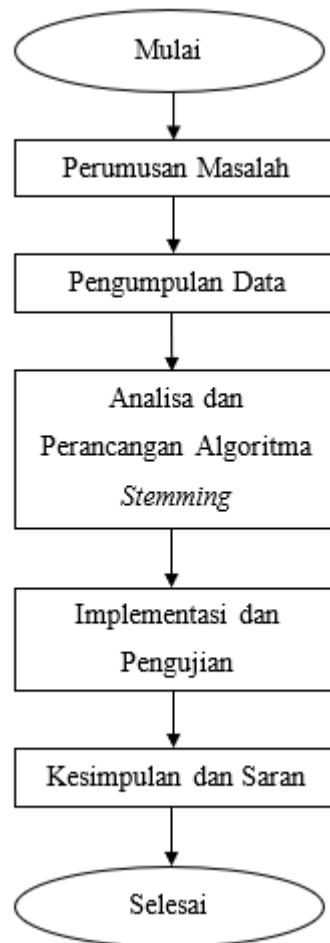
Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB III

METODOLOGI PENELITIAN

Metodologi penelitian menjelaskan tentang langkah-langkah yang ditempuh saat melakukan penelitian untuk menjawab pertanyaan dari rumusan masalah pada penelitian. Adapun langkah-langkah penelitian dapat dilihat pada Gambar 3.1 sebagai berikut:



Gambar 3. 1 Tahap Metode Penelitian

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

3.1. Perumusan Masalah

Pada langkah ini yang dilakukan adalah memahami permasalahan yang menjadi fokus penelitian. Ruang lingkup dan jenis data yang dibutuhkan pada penelitian juga ditentukan di tahapan ini. Adapun perumusan masalah pada penelitian ini adalah bagaimana membuat algoritma *stemming* penentuan kata dasar dari bahasa Minang dengan menerapkan aturan pembentukan bahasa Minang.

3.2. Pengumpulan Data

Pada tahap ini pengumpulan data diperoleh dari Kamus Bahasa Daerah Minang dan dari beberapa lirik lagu Bahasa Minang. Setelah kata dasar dikumpulkan, kemudian diinputkan ke dalam *database* untuk dijadikan data elektronik.

Pada penelitian ini, digunakan Buku Tata Bahasa dan Lirik Lagu Bahasa Minang atau teks Bahasa Minang untuk mencari aturan pembentukan kata berimbuhan.

3.3. Analisa dan Perancangan Algoritma *Stemming*

Pada langkah ini yang dilakukan adalah analisa data dan analisa pembentukan algoritma. Selengkapnya pada pembahasan dibawah ini:

3.3.1. Analisa Kebutuhan Data

Dalam menganalisa proses *stemming* bahasa Minang data yang digunakan pada penelitian ini berupa kamus lengkap bahasa Minang-Indonesia (Departemen Pendidikan dan Kebudayaan, 1985) dan morfologi bahasa Minang oleh (Departemen Pendidikan dan Kebudayaan, 1979). Data ini kemudian akan dimasukkan ke dalam database dengan cara mengetik ulang isi buku. Merumuskan aturan-aturan dalam pembentukan kata imbuhan dalam bahasa Minang.

3.3.2. Analisa Proses *Stemming*

Tahapan ini dilakukan untuk menemukan kata dasar dari kata berimbuhan pada bahasa Minang sesuai dengan aturan tata bahasa Minang. Kata dasar yang diperoleh dari proses *stemming* sesuai dengan kata dasar bahasa Minang yang berpedoman pada kamus bahasa Minang. Proses *stemming* dapat dilakukan setelah membuat algoritma untuk *stemming* bahasa Minang sesuai dengan aturan tata bahasa

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpulkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Minang. Tahapan yang dilakukan pada proses *stemming* yaitu dengan mengumpulkan aturan kata imbuhan seperti prefiks, sufiks dan konfiks, sesuai dengan aturan tata bahasa Minang. Aturan tersebut kemudian akan digunakan sebagai berikut:

1. Menghilangkan imbuhan awalan (prefiks) yang melekat pada sebuah kata.
2. Menghilangkan imbuhan akhiran (sufiks) yang melekat pada sebuah kata.
3. Menghilangkan imbuhan awalan akhiran (konfiks) yang melekat pada sebuah kata.

3.3.3 Perancangan

Tahapan perancangan merupakan tahapan setelah analisa algoritma. Sistem yang akan dirancang yaitu algoritma *stemming* bahasa Minang berdasarkan analisa dari langkah-langkah sebelumnya. Berikut ini langkah-langkah yang akan dilakukan, antara lain:

- a. Perancangan *flowchart* digunakan untuk menggambarkan alur dari proses algoritma yang akan dibuat dan hubungan antara suatu proses secara mendetail.
- b. Perancangan pseudocode digunakan untuk mendeskripsikan algoritma *stemming* yang akan dibuat menggunakan konvensi struktural agar dapat dibaca oleh manusia.
- c. Perancangan database dilakukan untuk mendukung rancangan sistem dan menentukan isi dari pengaturan data yang dibutuhkan. Berupa tabel yang dilengkapi dengan atribut kebutuhan algoritma *stemming* Bahasa Minang.
- d. Perancangan *interface* dibuat untuk penghubung antara pengguna dengan komputer. Dengan tujuan menciptakan tampilan yang mudah untuk dimengerti pengguna.

3.4. Implementasi dan Pengujian

Implementasi adalah penerapan dari hasil perancangan yang dilakukan sebelumnya sementara pengujian dilakukan dengan tujuan mengukur akurasi algoritma penentuan kata dasar dari bahasa Minang.

3.4.1. Implementasi

Untuk pengimplementasian sistem menggunakan bahasa pemrograman dan untuk perangkat yang akan digunakan adalah sebagai berikut:

Hak Cipta Diindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

© Hak cipta milik UIN Suska Riau

1. Perangkat Keras (*Hardware*)

Processor	: Intel® Core i5
RAM	: 4 GB
HDD	: 1 TB

2. Perangkat Lunak (*Software*)

Sistem Operasi	: Windows 10
Bahasa Pemrograman	: Python
DBMS	: PostgreSQL
Browser	: Google Chrome
Framework	: Django

3.4.2. Pengujian

Untuk pengujian algoritma dilakukan dengan tujuan apakah algoritma yang dibuat sudah sesuai berdasarkan perancangan yang dibuat. Pengujian menggunakan *white box* dengan akurasi menyiapkan kata berimbuhan kemudian diberi label manual kata dasar dan bandingkan dengan hasil algoritma. Sedangkan untuk pengujian data uji dilakukan dengan cara validasi. Validasi dilakukan oleh validator yang bernama Riky Mairizon, S.Sn. Beliau adalah pembina sanggar seni yang ada di Kota Payakumbuh-Sumatra Barat.

3.5. Kesimpulan dan Saran

Tahapan kesimpulan dan saran adalah tahapan terakhir dari penelitian. Pada kesimpulan berisikan hasil dari penelitian, akurasi dari algoritma penentuan kata dasar dari bahasa Minang. Selanjutnya pada bagian saran berisikan saran untuk pembaca agar dapat melakukan pengembangan dari penelitian yang sudah dilakukan.

UIN SUSKA RIAU

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

BAB VI

PENUTUP

6.1. Kesimpulan

Berdasarkan hasil penelitian yang telah dilakukan sebelumnya, maka dapat ditarik kesimpulan sebagai berikut:

1. Algoritma *stemming* bahasa Minang berhasil dilakukan berdasarkan aturan tata bahasa Minang.
2. Algoritma *stemming* yang telah dibuat berhasil mendapatkan tingkat akurasi sebesar 100% dari 345 kata uji dengan 345 kata uji yang benar.

6.2. Saran

Dari hasil yang telah didapatkan untuk pengembangan penelitian ini pada masa selanjutnya, maka diberikan beberapa saran sebagai berikut:

1. Penelitian Algoritma *stemming* bahasa Minang ini dapat dikembangkan untuk proses Normalisasi, Disambiguitas dan bahasa natural lainnya yang berhubungan dibidang NLP.
2. Untuk penelitian selanjutnya diperlukan penambahan kata dasar yang didapat tidak hanya dari lirik lagu sehingga data kata yang diperoleh akan lebih banyak.

DAFTAR PUSTAKA

- Arifin, D. S. (1990). *Kamus Tata Bahasa Lengkap Dengan Contoh*. Padang: Angkasa Raya.
- Ayub, Asni., Husin, N., Muhardi, Hakim Usman, A., Yasin, A., (1993). *Tata Bahasa Minangkabau*. Jakarta.
- BPS. (2010). *Kewarganegaraan, Suku Bangsa, Agama, Dan Bahasa Sehari-hari Penduduk Indonesia*.
- Burhanuddin, E., Syamsarul, Nasri, D., Arriyanti, Diana, Dewi, F., & Krisnawati. (2012). *Kamus Bahasa Minangkabau-Indonesia Balai Bahasa Padang Edisi Kedua*.
- James Suciadi Industri, F. T. (2004). Studi Analisis Metode-Metode Parsing dan Interpretasi Semantik Pada Natural Language Processing, 13–22.
- Manase Sahat H Simarankir. (2017). Studi Perbandingan Algoritma - Algoritma Stemming Untuk Dokumen Teks Bahasa Indonesia, 1(1), 40–46.
- Mar'atus Madia. (2016). Stemming Bahasa Jawa Untuk Mencari Akar Kata Dalam Bahasa Jawa Dengan Aturan Analisis Kontrasif Afiksasi Verba.
- Purwoko, A., & Pascasarjana, S. (2011). Berbasis kamus untuk dokumen berbahasa sunda.
- Rahmali, M., dkk (1985). *Kamus Bahasa Minangkabau-Indonesia*. Jakarta.

LAMPIRAN A

Tabel A. 1 Kata Uji Algoritma Stemming Bahasa Minang

No	Kata Imbuhan	Kata Dasar	Hasil Mesin	Hasil
1	Ambiakan	ambiak	ambiak	Benar
2	Babadan	badan	badan	Benar
3	Babaju	baju	baju	Benar
4	Babaleh	baleh	baleh	Benar
5	Babaliaklah	baliak	baliak	Benar
6	Babelok	belok	belok	Benar
7	Babendi	bendi	bendi	Benar
8	Babimbiang	bimbiang	bimbiang	Benar
9	Babuah	buah	buah	Benar
10	Babuek	buek	buek	Benar
11	Babungo	bungo	bungo	Benar
12	Bacarai	carai	carai	Benar
13	Bacinto	cinto	cinto	Benar
14	Badaki	daki	daki	Benar
15	Badarai	darai	darai	Benar
16	Badaun	daun	daun	Benar
17	Badawaik	dawaik	dawaik	Benar
18	Baduo	duo	duo	Benar
19	Bagageh-gageh	gageh	gageh	Benar
20	Bagalombang	galombang	galombang	Benar
21	Bagarah	garah	garah	Benar
22	Bagotong royong	gotong royong	gotong royong	Benar
23	Bairik	irik	irik	Benar
24	Bajalan	jalan	jalan	Benar
25	Bajapuik	japuik	japuik	Benar
26	Bajauah-jauahan	jauah	jauah	Benar

- Hak Cipta Dilindungi Undang-Undang**
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
 2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

27	Bajauahan	jauah	jauah	Benar
28	Bakarajo	karajo	karajo	Benar
29	Bakareh	kareh	kareh	Benar
30	Bakasiah	kasiah	kasiah	Benar
31	Bakirim	kirim	kirim	Benar
32	Bakuah	kuah	kuah	Benar
33	Bakuantam	antam	antam	Benar
34	Bakuliliang	kuliliang	kuliliang	Benar
35	Bakumpua	kumpua	kumpua	Benar
36	Bakupiah	kupiah	kupiah	Benar
37	Bakutinju	tinju	tinju	Benar
38	Baladang	ladang	ladang	Benar
39	Balangang	langang	langang	Benar
40	Balangang-langang	langang	langang	Benar
41	Baliku	liku	liku	Benar
42	Balinduang	linduang	linduang	Benar
43	Bamacam	macam	macam	Benar
44	Bamain	main	main	Benar
45	Baminantu	minantu	minantu	Benar
46	Banamo	namo	namo	Benar
47	Banyanyi	nyanyi	nyanyi	Benar
48	Baonggok	onggok	onggok	Benar
49	Bapacik	pacik	pacik	Benar
50	Bapaga	paga	paga	Benar
51	Bapaliang	paliang	paling	Benar
52	Bapaneh	pane	pane	Benar
53	Bapasan	pasan	pasan	Benar
54	Bapasulik	sulik	sulik	Benar
55	Bapatajam	tajam	tajam	Benar

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

56	Bapitih	pitih	pitih	Benar
57	Baputa	puta	puta	Benar
58	Baputiah	putiah	putiah	Benar
59	Baraka	aka	aka	Benar
60	Barakik-rakik	rakik	rakik	Benar
61	Barambang	rambang	rambang	Benar
62	Barampek	ampek	ampek	Benar
63	Baranak	anak	anak	Benar
64	Baranang-ranang	ranang	ranang	Benar
65	Baranti	anti	anti	Benar
66	Barayun	ayun	ayun	Benar
67	Barubah	ubah	ubah	Benar
68	Barumah	rumah	rumah	Benar
69	Basabab	sabab	sabab	Benar
70	Basajuak	sajuak	sajuak	Benar
71	Basakik	sakik	sakik	Benar
72	Basamo	samo	samo	Benar
73	Basampan	sampan	sampan	Benar
74	Basanang	sanang	sanang	Benar
75	Basandiang	sandiang	sandiang	Benar
76	Basayang	sayang	sayang	Benar
77	Basikareh	kareh	kareh	Benar
78	Basilambek	lambek	lambek	Benar
79	Basimpang	simpang	simpang	Benar
80	Basimpuah	simpuah	simpuah	Benar
81	Basitagang	tagang	tagang	Benar
82	Basubang	subang	subang	Benar
83	Basudahi	sudah	sudah	Benar
84	Basuliang-suliang	suliang	suliang	Benar

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

85	Basusah	susah	susah	Benar
86	Batambah	tambah	tambah	Benar
87	Batambuah	tambuah	tambuah	Benar
88	Batamu	tamu	tamu	Benar
89	Batanak	tanak	tanak	Benar
90	Batanam	tanam	tanam	Benar
91	Bataun	taun	taun	Benar
92	Batenggang	tenggang	tenggang	Benar
93	Batimbo	timbo	timbo	Benar
94	Batinggi-tinggian	tinggi	tinggi	Benar
95	Batingkah	tingkah	tingkah	Benar
96	Batuka	tuka	tuka	Benar
97	Batungkek	tungkek	tungkek	Benar
98	Baungkai	ungkai	ungkai	Benar
99	Bausaho	usaho	usaho	Benar
100	Bisiakkan	bisiak	bisiak	Benar
101	Bisuaklah	bisuak	bisuak	Benar
102	Bukannyo	bukan	bukan	Benar
103	Buuakan	buua	buua	Benar
104	Candonyo	cando	cando	Benar
105	Carikan	cari	cari	Benar
106	Diampaian	ampai	ampai	Benar
107	Dibaok	baok	baok	Benar
108	Dibasahi	basah	basah	Benar
109	Dibatangnyo	batang	batang	Benar
110	Dibuek	buek	buek	Benar
111	Dicamehkan	cameh	cameh	Benar
112	Digayuang	gayuang	gayuang	Benar
113	Diguntiang	guntiang	guntiang	Benar

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

114	Diharokkan	harok	harok	Benar
115	Dikacaunyo	kacau	kacau	Benar
116	Dikambangkan	kambang	kambang	Benar
117	Dilacuik	lacuik	lacuik	Benar
118	Dilambuik	lambuik	lambuik	Benar
119	Diliek	liek	liek	Benar
120	Dilingkuang	lingkuang	lingkuang	Benar
121	Dipakaikan	pakai	pakai	Benar
122	Dipasarabuikan	rabuik	rabuik	Benar
123	Dipasuntiang	suntiang	suntiang	Benar
124	Disabuik	sabuik	sabuik	Benar
125	Disangko	sangko	sangko	Benar
126	Diubek	ubek	ubek	Benar
127	Dutoi	duto	duto	Benar
128	Ganggaman	ganggam	ganggam	Benar
129	Garami	garam	garam	Benar
130	Guloi	gulo	gulo	Benar
131	Gunjiangan	gunjiang	gunjiang	Benar
132	Hapuihkan	hapuih	hapuih	Benar
133	Inkari	ingkar	ingkar	Benar
134	Janguaklah	janguak	janguak	Benar
135	Japuikkan	japuik	japuik	Benar
136	Jauahi	jauah	jauah	Benar
137	Kaampekk	ampekk	ampekk	Benar
138	Kabakeh	bakeh	bakeh	Benar
139	Kabalakang	balakang	balakang	Benar
140	Kadalam	dalam	dalam	Benar
141	Kagantino	ganti	ganti	Benar
142	Kakasiak	kasiak	kasiak	Benar

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

143	Kalua	lua	lua	Benar
144	Kambanglah	kambang	kambang	Benar
145	Kamuko	muko	muko	Benar
146	Kapadehan	padeh	padeh	Benar
147	Kapanuahan	panuah	panuah	Benar
148	Karauang	rauang	rauang	Benar
149	Kasanang	sanang	sanang	Benar
150	Kasudahannyo	sudah	sudah	Benar
151	Kusuiknyo	kusuik	kusik	Benar
152	Lagunyo	lagu	lagu	Benar
153	Lakehlah	lakeh	lakeh	Benar
154	Maawai	awai	awai	Benar
155	Mahampai	hampai	hampai	Benar
156	Maikek	ikek	ikek	Benar
157	Maimbau	imbau	imbau	Benar
158	Maimbau-imbau	imbau	imbau	Benar
159	Malabiah	labiah	labiah	Benar
160	Malangkah	langkah	langkah	Benar
161	Malapeh	lapeh	lapeh	Benar
162	Malengah	lengah	lengah	Benar
163	Malenggokkan	lenggok	lenggok	Benar
164	Malipek	lipek	lipek	Benar
165	Malukoi	luko	luko	Benar
166	Malupoan	lupo	lupo	Benar
167	Mamadu	madu	madu	Benar
168	mamaek	paek	paek	Benar
169	Mamakan	makan	makan	Benar
170	Mamaliharo	paliharo	paliharo	Benar
171	Mamandang	pandang	pandang	Benar

172	Mamanjek	panjek	panjek	Benar
173	Mamatahan	patah	patah	Benar
174	Mamatiak	patiak	patiak	Benar
175	Mambakeh	bakeh	bakeh	Benar
176	Mambaok	baok	baok	Benar
177	Mambasahi	basah	basah	Benar
178	Mambatu	batu	batu	Benar
179	Mamingik	bingik	bingik	Benar
180	Mambinguangkan	binguang	binguang	Benar
181	Mambunuah	bunuah	bunuah	Benar
182	Mampabudak	budak	budak	Benar
183	Mampacaliakan	caliak	caliak	Benar
184	Mampaduokan	duo	duo	Benar
185	Mamparancak	rancak	rancak	Benar
186	Mamutuih	putuih	putuih	Benar
187	Manabuih	tabuih	tabuih	Benar
188	Mananguang	tanguang	tanguang	Benar
189	Manangih	tangih	tangih	Benar
190	Manantikan	nanti	nanti	Benar
191	Manapek	tapek	takep	Benar
192	Manarawang	tarawang	tarawang	Benar
193	Manari	tari	tari	Benar
194	Mancabuik	cabuik	cabuik	Benar
195	Mancaliak	caliak	caliak	Benar
196	Mancari	cari	cari	Benar
197	Mancintoi	cinto	cinto	Benar
198	Mancogok	cogok	cogok	Benar
199	Mancubo	cubo	cubo	Benar
200	Mandakek	dakek	dakek	Benar

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

201	Mandaki	daki	daki	Benar
202	Mandalang	dalang	dalang	Benar
203	Mandanga	danga	danga	Benar
204	Mandayo	dayo	dayo	Benar
205	Mandukuang	dukuang	dukuang	Benar
206	Manembak	tembak	tembak	Benar
207	Mangadu	kadu	kadu	Benar
208	Mangakeh	kakeh	kakeh	Benar
209	Mangandakan	kandak	kandak	Benar
210	Mangantuak	kantuak	kantuak	Benar
211	Mangarek	karek	karek	Benar
212	Mangarokkan	arok	arok	Benar
213	Manggali	gali	gali	Benar
214	Manggumam	gumam	gumam	Benar
215	Mangukua	kukua	kukua	Benar
216	Manjalang	jalang	jalang	Benar
217	Manjawek	jawek	jawek	Benar
218	Manompang	tompang	tompang	Benar
219	Manonton	nonton	nonton	Benar
220	Manuai	tuai	tuai	Benar
221	Manulih	tulih	tulih	Benar
222	Manumbuak	tumbuak	tumbuak	Benar
223	Manurun	turun	turun	Benar
224	Manusuak	tusuak	tusuak	Benar
225	Manutuik	tutuik	tutuik	Benar
226	Manyambah	sambah	sambah	Benar
227	Manyambuang	sambuang	sambuang	Benar
228	Manyangko	sangko	sangko	Benar
229	Manyapo	sapo	sapo	Benar

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

230	Manyapu	sapu	sapu	Benar
231	Manyibak	sibak	sibak	Benar
232	Manyirami	siram	siram	Benar
233	Manyongsong	songsong	songsong	Benar
234	Manyuruah	suruah	suruah	Benar
235	Maramuak	ramuak	ramuak	Benar
236	Marandang	randang	randang	Benar
237	Marangguik	rangguik	rangguik	Benar
238	Marapek	rapek	rapek	Benar
239	Marendo	rendo	rendo	Benar
240	Marindu	rindu	rindu	Benar
241	Muluiknyo	muluik	muluik	Benar
242	Paarek	arek	arek	Benar
243	Pabudak	budak	budak	Benar
244	Pabuek	buek	buek	Benar
245	Pangunjiangan	gunjiang	gunjiang	Benar
246	Pailah	pai	pai	Benar
247	Pajalanan	jalan	jalan	Benar
248	Paketekan	ketek	ketek	Benar
249	Paladang	ladang	ladang	Benar
250	Palapeh	lapeh	lapeh	Benar
251	Pambaluik	baluik	baluik	Benar
252	Panakuik	takuik	takuik	Benar
253	Panapi	tapi	tapi	Benar
254	Pandakian	daki	daki	Benar
255	Pandatang	datang	datang	Benar
256	Pandayuang	dayuang	dayuang	Benar
257	Pandinginan	dingin	dingin	Benar
258	Panduduaknyo	panduduak	panduduak	Benar

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumpukan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

259	Panjilek	jilek	jilek	Benar
260	Panokok	tokok	tokok	Benar
261	Panutuik	tutuik	tutuik	Benar
262	Panyambuung	sambuung	sambuung	Benar
263	Paota	ota	ota	Benar
264	Parami	rami	rami	Benar
265	Parancah	rancah	rancah	Benar
266	Parubahan	ubah	ubah	Benar
267	Pasangkan	pasang	pasang	Benar
268	Patenggangan	tenggang	tenggang	Benar
269	Paubek	ubek	ubek	Benar
270	Pauruik	uruik	uruik	Benar
271	Putuihlah	putuih	putuih	Benar
272	Racuni	racun	racun	Benar
273	Ramilah	rami	rami	Benar
274	Rasonyo	raso	raso	Benar
275	Sadaguik	daguik	daguik	Benar
276	Sagadang	gadang	gadang	Benar
277	Sagantang	gantang	gantang	Benar
278	Sahiduik	hiduik	hiduik	Benar
279	Sairiang	iriang	iriang	Benar
280	Sakabek	kabek	kabek	Benar
281	Sakali	sakali	sakali	Benar
282	Sakampuang	kampuang	kampuang	Benar
283	Saksinyo	saksi	saksi	Benar
284	Sakuliliang	kuliliang	kuliliang	Benar
285	Samakin	makin	makin	Benar
286	Samalam	malam	malam	Benar
287	Samati	mati	mati	Benar

288	Sanduaklah	sanduak	sanduak	Benar
289	Sanjuangan	sanjuang	sanjuang	Benar
290	Sapaduo	duo	duo	Benar
291	Sapamakan	makan	makan	Benar
292	Sapambaran	bari	bari	Benar
293	Sapamberang	berang	berang	Benar
294	Sapantun	pantun	pantun	Benar
295	Sapanuah	panuah	panuah	Benar
296	Sapanyakik	panyakik	panyakik	Benar
297	Sapiriang	piriang	piriang	Benar
298	Sarato	rato	rato	Benar
299	Sarumah	rumah	rumah	Benar
300	Sarumpun	rumpun	rumpun	Benar
301	Satahun	tahun	tahun	Benar
302	Singgahlah	singguh	singguh	Benar
303	Tabanek	banek	banek	Benar
304	Tabangkan	tabang	tabang	Benar
305	Tacangang	cangang	cangang	Benar
306	Tadorong	dorong	dorong	Benar
307	Tagamang	gamang	gamang	Benar
308	Tagantuang	gantuang	gantuang	Benar
309	Tahampeh	hampeh	hampeh	Benar
310	Tajamua	jamua	jamua	Benar
311	Takambang	kambang	kambang	Benar
312	Takareh	kareh	kareh	Benar
313	Takayo	kayo	kayo	Benar
314	Takicuah	kicuah	kicuah	Benar
315	Talatak	latak	latak	Benar
316	Talinteh	linteh	linteh	Benar

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

317	Taluko	luko	luko	Benar
318	Tamakan	makan	makan	Benar
319	Tamanuang	manuang	manuang	Benar
320	Tanami	tanam	tanam	Benar
321	Tandonyo	tando	tando	Benar
322	Tanguangkan	tanguang	tanguang	Benar
323	Tapacik	pacik	pacik	Benar
324	Tapadalam	dalam	dalam	Benar
325	Tapaik	paik	paik	Benar
326	Tapakok	pakok	pakok	Benar
327	Tapandai	pandai	pandai	Benar
328	Tapanggang	panggang	panggang	Benar
329	Tapisah	pisah	pisah	Benar
330	Taputih	putiah	putiah	Benar
331	Taputih	putih	putih	Benar
332	Tarandam	randam	randam	Benar
333	Taramuak	ramuak	ramuak	Benar
334	Tasakik	sakik	sakik	Benar
335	Tasanguik	sanguik	sanguik	Benar
336	Tasintak	sintak	sintak	Benar
337	Tasirah	sirah	sirah	Benar
338	Tasisiah	sisiah	sisiah	Benar
339	Tatagah	tagah	tagah	Benar
340	Tatumpah	tumpah	tumpah	Benar
341	Tibonyo	tibo	tibo	Benar
342	Tinggalah	tingga	tingga	Benar
343	Turuikkan	turuik	turuik	Benar
344	Ubekkan	ubek	ubek	Benar
345	Usahlah	usah	usah	Benar

Hak Cipta Dilindungi Undang-Undang

1. Diarangi mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LAMPIRAN B

Surat Persetujuan Validator

Surat Persetujuan

Saya yang bertanda tangan dibawah ini:

Nama : Riky Mairizon S.Sn
Tempat / Tanggal Lahir : Payakumbuh / 6 Mei 1987
E-mail : rikymairizon1987@gmail.com
No. HP : 0821-7027-7407
Pendidikan : S1 Seni Teater ISI PADANG PANJANG
Aktivitas : Aktif Membina Sanggar
Alamat : Payobasung Kota Payakumbuh

Dengan ini menyatakan bahwa saya bersedia untuk menjadi Validator dari penelitian:

Nama : Dewita Putri
NIM : 11351201047
Judul Penelitian : Algoritma *Stemming* Bahasa Minang Berbasis Aturan Tata Bahasa
Instansi : Universitas Islam Negeri Sultan Syarif Kasim - Riau

Payakumbuh, 20 Desember 2020

Hormat Saya,



(Riky Mairizon S.Sn)

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LAMPIRAN C

Surat Validasi Data

Lembar Validasi Data *Stemming*

Saya yang bertanda tangan dibawah ini:

Nama : Ricky Mairizon S.Sn
Tempat / Tanggal Lahir : Payakumbuh / 6 Mei 1987
E-mail : rikymairizon1987@gmail.com
No. HP : 0821-7027-7407
Pendidikan : S1 Seni Teater ISI PADANG PANJANG
Aktivitas : Aktif Membina Sanggar
Alamat : Payobasung Kota Payakumbuh

Dengan ini menyatakan bahwa hasil data penelitian yang berjudul “Algoritma *Stemming* Bahasa Minang Berbasis Aturan Tata Bahasa” telah sesuai dengan kata dasar yang sebenarnya dari kata berimbuhan bahasa Minang yang diberikan.

Demikian lembar validasi ini dibuat untuk digunakan dalam melengkapi Laporan Tugas Akhir.

Payakumbuh, 20 Desember 2020

Hormat Saya,



(Riky Mairizon S.Sn)

Hak Cipta Dilindungi Undang-Undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

LAMPIRAN D

Data Validator

DATA VALIDATOR



Nama	: Riky Mairizon S.Sn
Tempat / Tanggal Lahir	: Payakumbuh / 6 Mei 1987
Jenis Kelamin	: Laki-laki
Alamat	: Jln. Sultan Hasanuddin RT 007/ RW 001 Kelurahan Padang Tongah Balai Nan Duo Kec. Payakumbuh Barat
Agama	: Islam
Status Perkawinan	: Belum Kawin
Pekerjaan	: Seniman

UIN SUSKA RIAU

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber:
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah.
 - b. Pengutipan tidak merugikan kepentingan yang wajar UIN Suska Riau.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin UIN Suska Riau.

DAFTAR RIWAYAT HIDUP



Nama Lengkap : Dewita Putri
 Tempat/Tanggal Lahir : Pekanbaru / 15 Januari 1996
 Nama Ayah : Mofrizal
 Nama Ibu : Zulbaida
 Anak Ke- : 2 (Dua)
 Jumlah Saudara : 4 (Empat)
 Alamat : Jl. Riau Gg. Angkasa No. 14-D
 Pekanbaru, Riau
 E-mail : dewita.putri@students.uin-suska.ac.id

PENDIDIKAN

- ❖ Tahun 2001-2007 : SD Negeri 001 Tampan
- ❖ Tahun 2007-2010 : SMP Negeri 12 Pekanbaru
- ❖ Tahun 2010-2013 : SMK Negeri 1 Pekanbaru
- ❖ Tahun 2013-2021 : Universitas Islam Negeri Sultan Syarif Kasim Riau
Jurusan Teknik Informatika

UIN SUSKA RIAU